METHODS OF EDITING CLOUD AND ATMOSPHERIC LAYER AFFECTED PIXELS FROM SATELLITE DATA

Principal Investigator: Paul R. Nixon
Co-Investigators: Craig L. Wiegand
Arthur J. Richardson
Michael P. Johnson

Agricultural Research Service
U.S. Department of Agriculture
P. O. Box 267
Weslaco, TX 78596

November 1982

Type III Final Report for Period
December 5, 1980 To June 5, 1982

Prepared for
GODDARD SPACE FLIGHT CENTER
Greenbelt, MD 20771

Table of Contents

1.0 INTRODUCTION

The purpose of this study was to develop practical methods of computer screening cloud-contaminated pixels from data of various satellite systems. Removal of such pixels prior to interpretation would avoid erroneous conclusions about Earth surface conditions such as extent and severity of, for example, droughts and freezes and their probable economic impact due to crop losses.

The chapters that follow address the three topics of our proposal for this investigation. An additional chapter illustrates how the findings reported here can be used in editing cloud and atmospheric layer affected pixels from satellite data. Specifically:

In Chapter 2 examples are given of the location of clouds and representative landscape features in HCMM spectral space of reflectance (VIS) vs emisssion (IR). Methods of screening out cloud affected HCMM pixels are discussed.

In Chapter 3 the character of subvisible absorbing-emitting atmospheric layers in HCMM data are considered. These layers we refer to loosely as subvisible cirrus (SCi). Radiosonde soundings are examined in relation to the presence of SCi.

In Chapter 4 the statistical characteristics of multispectral meteorological sattellite data in clear and SCi affected areas are discussed. Examples in TIROS-N and NOAA-7 data from several states and Mexico are presented.

In Chapter 5 the VIS-IR cluster screening method for removing clouds is applied to a 262,144 pixel HCMM scene from south Texas and northeast Mexico. The SCi that remain after cluster screening are edited out by applying a statistically determined IR limit.

2.0 SCREENING CLOUDS IN HCMM DATA

2.1 Methods of screening

The cloud screening methods mentioned below have been tried with 8 daytime HCMM scenes from south Texas (Wiegand, et al., 1980). The clouds were mainly of the "popcorn" cumulus type and covered from 3 to 26% of the 100,000 pixel test area of the Lower Rio Grande Valley of Texas.

This chapter deals with screening clouds that are highly reflective in the visible wavelengths (relative to terrestrial features except snow) from satellite data. Discussion of screening subvisible radiation-absorbing-emitting atmospheric layers is reserved for later chapters. These atmospheric layers that are readily discernible in emissive images, but are not apparent in reflective images we loosely refer to as subvisible cirrus (Sci).

2.1.1 Reflective (VIS, 0.55 to 1.1 um) screening

Screening reflective (VIS) data above a threshold value removed many of the clouds but there were areas of uncertainty. If the threshold was set low enough to detect thin clouds and pixels partly contaminated by clouds, some reflective land areas such as fallow soil and dune areas were excluded. Also, the reflective criterion did not remove the cloud shadow.

2.1.2 Emissive (IR, 10.5 to 12.5 um) screening

Screening solely on the basis of emission was not feasible because water and low-altitude clouds tend to have overlapping temperature values.

2.1.3 Reflective/emissive ratio screening

The ratioing of reflective (VIS) to emissive (IR) pixel values was tried for screening daytime clouds. This required that the data from the two wavelengths be registered so they could be overlaid and ratioed pixel by pixel.

Screening for clouds by the reflective to emissive ratio was superior to either reflective or emissive screening alone. The ratio is useful because clouds are the brightest (highest digital count (DC)) scene component in the reflective channel and clouds are usually colder (lower digital count) than land features in the emissive data. Thus, the ratios, VIS/IR or IR/VIS, contrasted clouds very well versus other scene features for a given overpass date.

A weakness of the ratio method is that the division between clouds and terrestial features is not a fixed ratio but varies from scene to scene and with season.

2.1.4 Standard deviation screening

Tarpley (1979) described a process for eliminating cloud contaminated observations from GOES VISSR thermal data using standard deviations. He reasoned that the standard deviations for cloud-contaminated data would be more variable than noncloud contaminated data. However, data for water body-land interfaces such as for Falcon Reservoir, the Gulf of Mexico, and the bay (Laguna Madre) of the Lower Rio Grande Valley test area, also have high standard deviations; thus Tarpley's scheme would incorrectly classify these water body-land interface areas as clouds. Also, the subvisible cirrus (SCi) contaminated area does not yield a high standard deviation; thus, Tarpley's scheme would incorrectly identify the SCi contaminated areas as clear.

It is possible that the Great Plains area, where Tarpley's scheme was applied, had a dry enough atmosphere and a dearth of water bodies so that neither was a problem. However, there are many other places in the world where this may be a problem.

2.1.5 Moving-window screening

A cloud detection scheme was applied to emissive scenes that uses the ratio of the center pixel value (Xc) to the average pixel value ($\bar{X}w$) of a moving nxn pixel size window. Variations in Xc due to clouds are treated as high frequency image detail that is partially blocked by the low pass filter, $\bar{X}w$ data. Ratio values less than $Xc/\bar{X}w=0.92$ we found experimentally to be due to a partial cloud and the corresponding center pixel values were censored. The value 0.92 appears to be set by sensor NEΔT and the other noises.

This approach appears to distinguish between cloud values and surface and water body values satisfactorily. Unfortunately, it causes many valid pixels to be discarded in the vicinity of cloud-free and cloud-contaminated pixels in the nxn pixel window used as a filter. This may not be ojectionable where generalized results are desired. The method does not censor vapor contamination.

2.1.6 Cluster screening

In another cloud screening method, we took advantage of the fact that clouds are individualistic in scatter diagrams of reflective and emissive pixel values. Clouds have few locations where pixel values of both wavebands are the same as for the other scene features. That is, in two-dimensional spectral space, clouds do not cluster. Conversely, pixel values from Earth's surface tend to have numerous locations where both pixel values are the same as other locations. Thus, clusters are formed. The scatter in cloud data and the clustering of clear areas were used to separate data

with clouds from data for noncloudy areas.

Gray maps were prepared, in the initial HCMM study, after discarding all reflective-emissive combinations that did not recur at least 17 times in the test area of approximately 100,000 pixels. Combinations that recurred at least 17 times were considered to be land or water features. The gray maps of the study area, that were prepared on a computer line printer, represented various features or classes by the intensity (boldness) of the printed symbols.

It should be pointed out that the cut off percentage of 0.017 is not a fixed value but varies with the amount of cloud cover. A further consideration is that the cut-off percentage varies with the number of digital counts that represent the full VIS and IR ranges. TIROS-N series satellites have greater digital count ranges, hence their cut-off percentage is smaller than for HCMM.

In HCMM scenes of our 100,000 pixel Lower Rio Grande Valley test area uncontaminated land and water features have shared combinations of VIS and IR counts that occur with at least these frequencies.

Percent of scene cloud contaminated	Threshold percent of shared VIS vs IR combination for cloud-free feature
0	0.0213
5	0.0203
10	0.0192
15	0.0181
20	0.0171
25	0.0160
30	0.0149

The more cloud contaminated a scene is, the lower the incidence of clear pixels sharing the same VIS and IR digital count combination. As clouds increase, more contaminated pixels share VIS and IR combinations. This points to the desirability of excluding any large areas of cloud contamination from a scene at the start, before cluster screening is undertaken.

However, close examination of the gray maps prepared by the clustering criterion showed that some cloud shadows remained in the scene and certain land and water areas had been discarded.

2.1.7 Modified cluster screening

Of the above methods for screening clouds, clustering was the best. Most of the HCMM data processed at this location has been with a modification of the cluster method.

The modification was applied by examining a scatter diagram of reflective and emissive pixel coordinates. The analyst subjectively positioned a line (not necessarily straight) on the scatter diagram that separates cloud pixels from noncloud pixel coordinates. The line was then described mathematically for machine processing of the data. All pixels having reflective-emmissive coordinates that fall on the cloud side of the line were discarded.

A gray map prepared from the retained data was examined for undesirable data (such as cloud shadows) or inappropriate exclusion of valid features (such as highly reflective land areas). If the gray map was unsatisfactory, the line separating the clouds from valid data on the scatter diagram was adjusted and the procedure repeated.

2.2 Nature of HCMM data in spectral space.

In Figure 1, the approximately 100,000 HCMM Lower Rio Grande Valley area is seen in spectral space, as reflective (VIS) band vs emissive (IR) band digital counts. For this daytime scene, 3 Jul 78, the computer calculation of cloud cover at overpass was 18.8%. The number of pixels falling on the plotting positions in the figure are shown by these symbols:

.	1-4
/	5-10
V	11-16
$	17 or more

The main concentration of data represents land area which increases in temperature with distance inland. A smaller concentated area in the lower portion of the illustration represent water of the Gulf of Mexico and inland bodies of water. The latter form the warmer, more reflective cluster. The scattered data of low incidence in the upper and left portions of the graph represent clouds or cloud contaminated data.

Information about polar-orbiting satellites used in this study, overpass dates and times, locations of scene centers and sun zenith angles for the data of the figures of this report are given in Table 1.

In Figure 2 we show the retained data representing Earth surface features after modified cluster screening was applied to the 3 Jul 78 data. The analyst subjectively positioned a boundary on the scatter diagram in an attempt to separate cloud-contaminated pixels from pixels unaffected by clouds.

The distribution of occurrences (log scale) of VIS digital counts is shown in Figure 3 for all 100,000 pixels. Two peaks are

apparent in the VIS data where there were over 1000 occurrences for each digital count; in the 7 to 11 count range, representing water, and the 34 to 56 count range depicting land features. The land peak exceeded 5600 occurrences at a common VIS digital count, 42.

Distribution of VIS occurrences after screening out clouds is shown in Figure 4. On the insensitive log scale, the curves of Figures 3 and 4 are essentially identical up to 46 digital counts. The difference becomes prominent beyond that point; until no occurrences remain in the screened data above 72 digital counts.

Similarly, in Figures 5 and 6 we show the distribution of IR occurrences before and after cloud screening. The most common IR digital count value in the test area was 80 (24.9 C atmospherically corrected). It was common to both land and water and accounted for 2.6% of the pixels. In agreement with Figure 2, all IR digital counts of 65 or less have been removed by screening. The curves of Figures 5 and 6 are very similar above the 66 digital count, except for an additional dip between digital counts 82 and 103 in the screened data. Although not apparent with log scale, screening did cause some changes over the entire range from 66 to 112. The screened pixels in this range consisted of low altitude (warm) clouds and pixels that were mixtures of clouds and land features.

2.3 Location and migration of cloud and landscape features in spectral space.

The information presented here regarding the positions of cloud, land, and water features in spectral space (reflective vs. emissive) improves the understanding of HCMM data. Similar methods should apply to other satellite data.

The relationship on 3 Jul 78 of VIS (reflective) to IR (emissive) for the 100,000 pixels of our Lower Rio Grande Valley test area are shown in the upper portion of Figure 7. These are the same data as Figure 1 except that the symbols have been selected to emphasize the VIS-IR combinations of highest frequency.

The lightest plotting symbol represents VIS vs. IR combinations that experienced one or two pixels having a common combination. The boldness of the symbols increase with increasing number of pixels sharing the same position in spectral space. The centers of the land and water clusters are plotted with the darkest symbol which represents over 90 pixels sharing the same VIS-IR combination.

The scattered pixel locations in Figure 7 (upper) represent clouds or cloud contaminated pixels. They constitute 18.8% of the pixels on this day, 3 Jul 78.

There were apparently two major cloud types on 3 Jul 78. Pixels of one cloud type had low reflectivity--in the same range or

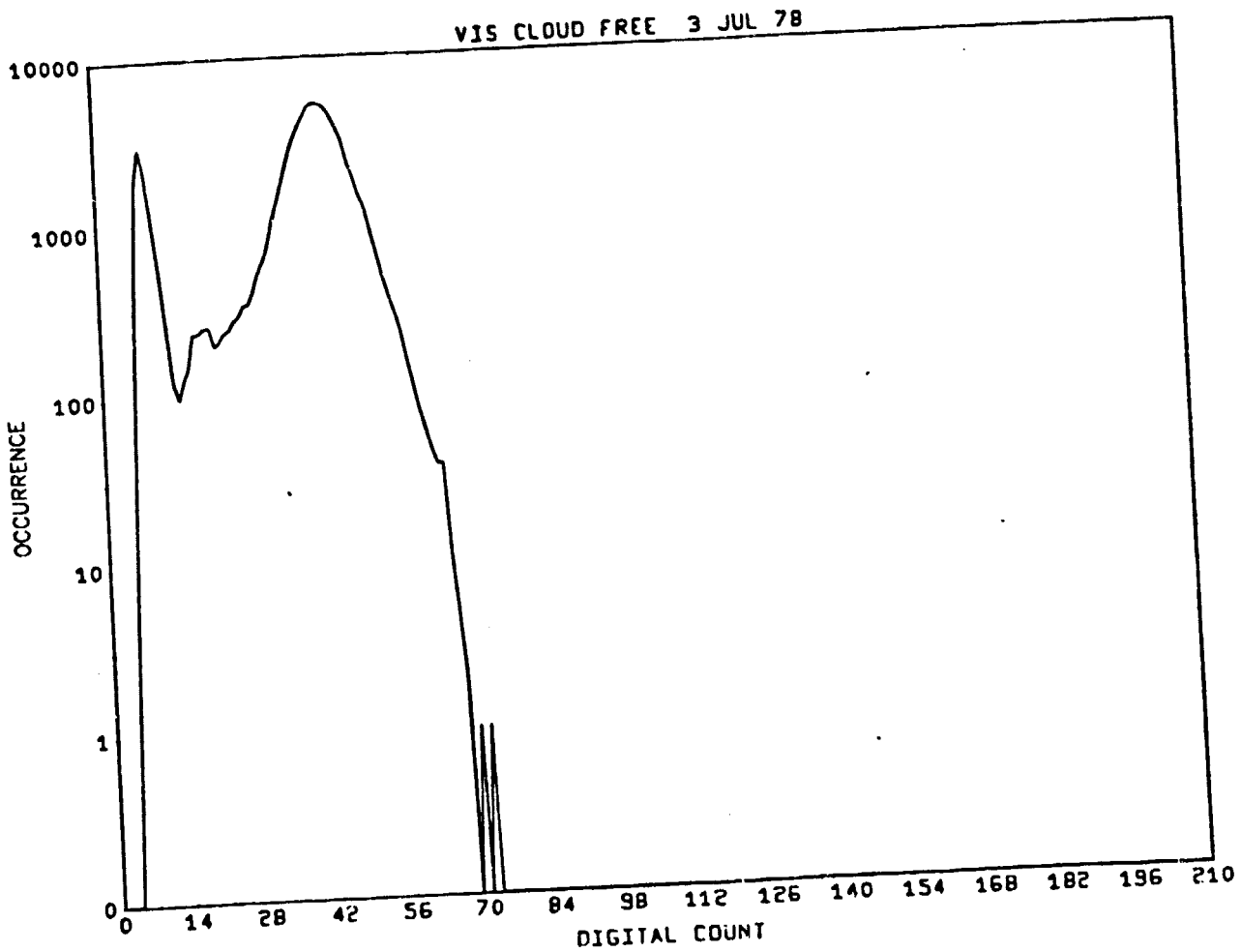

Figure 4. Distribution of remaining reflective (VIS) data in a 100,000 pixel Lower Rio Grande Valley of Texas test area on 3 Jul 78 after cloud screening by a modified cluster method.

The trend in Figure 7 (upper) is for the scattered points to increase in emissive counts and decrease in reflective counts as the contribution of land area to the pixels increases and the fraction of cloud contribution decreases.

Cumulus cloud shadows represented by S's plotted below the land cluster (Figure 7, lower) provided the lowest reflectances of the scene except for water bodies. Inland cloud shadows from high clouds indentified by H's, fell on the low edge of the land cluster but above the cumulus shadows. The H farthest to the left in the figure appears to have been under the influence of SCi.

Land. The most reflective land feature on 3 Jul was Padre Island dune sand, P's of Figure 7 (lower). The next most reflective land feature was the dryland sorghum, D's, which was mature or had been harvested by this date. The warmest feature was nonirrigated buffelgrass, G's, on an inland ranch.

Water. The tight cluster of W's in the lower portion of the figure and identification in Table 2 confirms that the concentrated small cluster was water of the Gulf of Mexico. Inland water bodies "A" were more reflective due to suspended solids, and slightly warmer on 3 Jul than the Gulf of Mexico.

Migration in Spectral Space. In Figure 8 we show that on 15 Aug 78 clouds were at a different location in spectral space than on 3 Jul. The tail to the left of the land cluster is missing. The SCi clouds represented by U's and V's were much warmer than the SCi on 3 Jul, being difficult to separate from the land cluster. Their VIS values were very similar to land features.

An essentially cloud free winter day, 7 Feb 79, is shown in spectral space by Figure 9. Cluster screening identified 0.2 percent clouds in the test area on this date. Obviously the land and water features were colder than they had been during the summer. Comparisons among figures show that the land cluster was more compact in winter than summer. This is reasonable considering the lower irradiation level at this time of year which would lessen the differences in land feature reflectances and temperatures.

3.0 ABSORBING - EMITTING ATMOSPHERIC LAYER EFFECTS IN HCMM DATA

3.1 Manifestation of the phenomenon

Absorbing-emitting atmospheric layer effects on satellite data are illustrated by HCMM VIS and IR scenes of 15 Aug 78 from a 262,144 pixel south Texas and northeast Mexico region (Figures 10 and 11). The scenes include the Lower Rio Grande Valley test area of our earlier HCMM study.

Visual comparison shows that cumulus clouds had the same pattern and extent in both wavelengths. However, there were additional patterns in the emissive scenes that had no counterparts in the reflective images. The more pronounced patterns seen in the emissive images were often observed in the reflective images, but the edges were very diffuse and effects were too subtle to be readily detected in automated processing of the digital reflective data. We attributed the patterns readily discernable in the emissive images, but not apparent in the reflective images, to absorbing-emitting atmospheric layers which we loosely refer to as subvisible cirrus (SCi).

Portions of the test area were under SCi layers in 6 of the 8 daytime, summertime HCMM south Texas scenes investigated. The wintertime 08 Nov 78 and 07 Feb 79 scenes that we studied were free of SCi layers.

3.2 Reflective and emissive profiles across subvisible cirrus streamer

Profiles across south Texas on latitude 26.47 N are illustrated in figure 12 for the HCMM daytime overpass of 15 Aug 78. The Gulf of Mexico is on the right side (east). Its low reflectance is apparent in the trace of Ch. 1 (VIS), and its water temperature of 22 C (uncorrected) is represented by the trace of Ch. 2 (IR). Padre Island, a coastal barrier island, is transected at approximately 97.3 W, showing a high reflectance and warmer temperature. Laguna Madre, between the island and the mainland, appears at 97.4 W. The mainland occupies the remainder of the figure with reflectance generally ranging from 14 to 22 percent, and surface temperatures increasing with distance inland (from right to left). An area of scattered cumulus clouds is shown between 97.6 and 97.9 W.

A subvisible streamer (SCi) present from 98.7 to 99.0 W is not apparent in the Ch. 1 trace of reflectance in Figure 12. In contrast the SCi caused a 10 C decrease in satellite-indicated surface temperature in the Ch. 2 trace. This SCi streamer being discussed is in the approximate center of Figure 11. (There was no rain so this was not caused by surface cooling resulting from wetting.) We interpret the figure to show that the effect was greatest in the center third of the SCi streamer, where it was most

dense, and tapered off to the edges. Similar results were obtained from transects along other latitudes in our test area during the summer season.

3.3 Comparison of clear and subvisible cirrus affected areas

The HCMM Ch. 1 and Ch. 2 digital characteristics of the SCi streamer mentioned above (and shown near the center of Figure 11) are compared in Figure 13 with an adjacent clear area of about equal size (498 vs. 496 pixels). The digital data of each channel were normalized to the clear area mean determined for that channel. The clear area mean was assigned a value of 100%.

The clear area data limits of ±2 standard deviations from the mean are shown by solid squares in Figure 13. The dotted lines connecting the clear area data serve only the purpose of guiding the eye between channels.

The SCi affected area data, normalized to mean clear area digital counts, are shown by bars. The diamond symbol represents the mean SCi value, and the extremes of the vertical bar give the ±2 standard deviation values.

It can be seen in Figure 13 that the SCi affected area is virtually indistinguishable from the clear area in the reflective VIS Ch. 1. The SCi data fell statistically within the clear area data. In contrast in the emissive IR Ch. 2 the SCi data were markedly different than the clear area data, being colder and showing less variability.

Comparisons are made in Figure 14 of another SCi area and its adjacent clear area. The ranges shown in this illustration are ±2 standard deviations. The SCi area was located at 26.9 N, 98.1 W and the HCMM information was collected on the same date as the previous example, 15 Aug 78. The SCi area appears in the IR scene of Figure 12 at a position about 1/3 from the top and 1/4 from the right edge.

In both examples (Figures 13 and 14) SCi is seen to be digitally different from clear areas in Ch. 2. This suggests that SCi can be screened from satellite data on the basis of statistical characteristics in the emissive channel. Pixels that vary from the scene mean by several standard deviations on the cold side are generally composed of clouds or SCi.

3.4 Upper air soundings in relation to subvisible cirrus

HCMM IR data of three cold nights in the Lower Rio Grande Valley test area were examined for indications of SCi. The presence of SCi is beneficial in retarding outgoing noctural radiation and hence slowing surface cooling, resulting in reduced freeze hazard on cold nights. Obviously SCi is undesirable if one wishes to

determine surface temperatures using satellite.

The nightime HCMM IR of the test area on 21 Oct 78 had about 6% cover which appeared to be SCi in nature, on 3 Jan 79 it was about 60% SCi covered (Wiegand et al. 1981, Nixon 1982), and 26 Feb 79 was clear.

The radiosonde observations that were made by the National Weather Service, Brownsville, Texas (within our test area) corresponding to these nights are presented in Figure 15. The data are from the 0600 hrs. soundings, except 3 Jan which is the 1800 hrs. sounding because the morning sounding is missing. These data collected following the passage of cold fronts suggest tendencies on 21 Oct and 3 Jan for ice crystal formation above the 600 mb. altitude. This is because converging temperature and dew point profiles indicate a moisture saturated condition, and either the existance of water droplet or ice crystals (below freezing), or their imminent formation.

It is concluded from these upper air observations and the varied location of SCi in HCMM scenes, that while the possibility of SCi may be inferred from radiosonde data, the certainty of the presence of SCi is not necessarily established. The weakness of using radiosonde measurements to determine the likehood of SCi in satellite scenes is that the soundings are made essentially at a point location and can not represent in detail the varying situation throughout the scene. An additional drawback is that upper air soundings are not customarily made at the time of the satellite overpass but at a time dictated by the world meteorological network.

4.0 ABSORBING - EMITTING ATMOSPHERIC LAYER EFFECTS IN METEORLOGICAL SATELLITE DATA

4.1 Subvisible cirrus in TIROS-N data

Examples of the statistical characteristics of subvisible cirrus (SCi) affected areas compared with adjacent clear areas, as observed in TIROS-N data, are given in Figures 16 and 17. These illustrations from the 6 May 79 daytime overpass represent locations in Texas and Mexico identified in the figures by the latitude and longitude of the center of the SCi.

These illustrations were prepared similarly to Figures 13 and 14. The figures are referenced to the mean digital value of the clear area, which was assigned 100%. The solid squares express the ±2 standard deviations of the clear area. The SCi data are represented by vertical bars, with a diamond depicting the mean value and the limits of the bar are the ±2 standard deviation values.

A difference between meteorological satellite and HCMM calibrations is that the increasing digital values in the emissive channels of meteorological satellite (TIROS-N Ch. 3 and Ch. 4; NOAA-7 Ch. 3, Ch. 4, and Ch. 5) represent colder temperatures, whereas increasing counts in HCMM represent warmer temperatures. Thus the colder temperatures of SCi areas with respect to clear areas plot above the clear area data in Figures 16 and 17.

It can be seen from the figures that TIROS-N reflective and emissive channels gave similar results as obtained with HCMM, namely that there was little difference between SCi and clear areas in Chs. 1 and 2, but large differences were present in Chs. 3 and 4.

A problem with TIROS-N has been noise in Ch. 3 data. Display of Ch. 3 on the image processing system showed the expected features of a mid-infrared (MIR) scene. However, close inspection revealed noise characterized by one or more "noise-free" scan lines followed by noisy lines, followed by noise-free lines, etc. By noise-free we mean free of unusual instrumentation errors. The noise-free and noisy lines did not generally prevail across the entire scene, but frequently changed from one class to the other, one or more times along the scan.

To test if Ch. 3 could be used in our study of SCi, we identified a noise-free 140 pixel area in the Gulf of Mexico for comparison with a nearby 450 pixel noisy area. The mean Ch. 3 counts were 625.2 and 620.0, respectively; not greatly different, considering the possible effect of Gulf currents. A similar comparison of adjacent noise-free and noisy Ch. 3 data in Texas (each less than 76 pixels) gave means of 456.8 and 459.0. These suggest that mean values for a given landscape feature are

representative of the feature whether they come from noise-free, mixed or noisy data, provided the sample is large enough. However, the standard deviations did vary markedly between noise-free and noisy data. They were 10.6 and 62.6 counts, respectively, for the Gulf of Mexico, and 18.8 and 78.0 for the land area. For the purposes of this study the Ch. 3 standard deviations were made "free" of the effects of unusual instrument error by factoring out the Ch. 3 IR noise on the basis of the comparisons just mentioned.

However a comparison of the TIROS-N results in Figures 16 and 17 with those of HCMM presented earlier (Figures 13 and 14) and NOAA-7 to follow, raises questions as to the efficacy of the procedure used to remove the effects of noise in TIROS-N Ch. 3 data. Nevertheless, in a figure to come (Figure 23) for NOAA-7 data from Kentucky, the deviation of SCi emissive data from the mean clear conditions are of the same magnitude as shown for Texas and Mexico in Figures 16 and 17.

4.2 Subvisible cirrus in NOAA-7 data

Examples of the statistical characteristic of SCi affected areas compared with adjacent clear areas as observed in NOAA-7 overpass of 23 Sep 81 are given in Figures 18, 19, and 20. These data are from Nebraska, South Dakota, and North Dakota. Data from the 29 Sep 81 NOAA-7 overpass of Iowa, Illinois, and Kentucky are given in Figures 21 through 23.

The figures were prepared similarly to Figures 13 through 17. The clear area mean was set to 100% in every case, and the ±2 standard deviations limits are shown.

As with the HCMM and TIROS-N data, the NOAA-7 show that SCi layers tend to be indistinguishable in the reflective (Chs. 1 and 2) but do have distinguishable statistical characteristics in the emissive (Chs. 3, 4, and 5).

In Figures 21 and 22 the same clear area was used for reference with the SCi layers in Illinois and Iowa. The Illinois SCi data separate more distictly than the Iowa SCi in Chs. 3, 4, and 5. Also note that the range of ±2 standard deviations is less in the dense Illinois SCi compared with the thinner Iowa layer. Relatively narrow emissive ranges have been observed with other dense SCi layers (Nixon 1982, Wiegand et al. 1981).

4.3 Statistical characteristics that aid editing

A review of Figures 13 and 14 and 16 through 24 suggests that the marked statistical differences between SCi affected areas and clear areas can be used in screening out SCi contamination from satellite data. If the clear and SCi data were both normally distributed, one could say that 95% of the respective data fall

within their ±2 standard deviation limits shown in the figures. This would mean only 2.5% of the data are outside of the limits on each side. Of course the data are only approximately normally distributed, however the illustrations are interpreted to show that in a number of cases the SCi and clear data are greatly separated and their respective standard deviation ranges are far from overlapping.

In several of the examples there is little choice among the various emissive channels as to the extent of statistical separation between SCi and clear areas. However, on the basis of our overall experience so far, Ch. 4 is generally the best choice for differentiating between SCi and clear in NOAA-7 data.

5.0 EXAMPLE OF EDITING CLOUD AND ATMOSPHERIC LAYER AFFECTED PIXELS FROM HCMM DATA

5.1 Approach used for editing

This study was undertaken with the knowledge that clouds and subvisible absorbing-emitting atmospheric layers affect satellite data, with some pixels being wholly or partially contaminated. Cloud affected pixels are undesirable in both reflective (VIS) and emissive (IR) data. The influence of subvisible cirrus (SCi) is harmful in IR data where it can cause errors of as much as 10 C in determining surface temperatures (Figure 12).

The findings of this report can be applied in the following way during data processing to rid daytime satellite data of cloud and SCi contaminated pixels. The procedure first removes cloud affected pixels, then the data that remain are purged of SCi contamination.

1. Exclude extensive solidly cloud covered areas from the data by defining a boundary within the scene between unusable and salvageable data.

2. Remove cloud affected pixel from salvageable data by applying a form of cluster screening in two-dimensional spectral space (VIS vs IR).

3. Examine thermal image of the scene to determine if SCi is present.

4. If SCi is present, determine the statistics of the IR data (mean and standard deviation of data that remain after cloud affected pixels have been removed).

5. Remove the coldest remaining pixels. They are generally SCi affected. They deviate from the IR mean (on the cold side) by several standard deviations, typically more than 2.5 standard deviations.

6. The pixels that remain after the two-stage screening are essentially free of cloud and SCi contamination.

5.2 An example of editing

What follows is an example of editing cloud and atmospheric layer affected pixels from daytime HCMM data. The example removes contaminated pixels from the 15 Aug 78 scene of south Texas and northeast Mexico that is shown in Figures 10 and 11. The procedures outlined above are followed in doing this.

Step 1, calling for the exclusion of solidly covered cloud areas, was not necessary because the scene did not include an extensive area of continuous clouds.

Step 2 is illustrated by Figures 24 and 25. The 262,144 pixels of the scene are shown plotted in spectral space, Ch. 1 VIS vs Ch. 2 IR, in Figure 24. The light-colored center of the cluster represents land features where each combination of VIS and IR values was repeated at least 120 times within the scene. In contrast the outer blue region represents VIS and IR combinations that were shared by 20 or fewer pixels. This low recurrence of shared values is characteristic of clouds.

On the basis of Figure 24, and information discussed earlier, pixels that shared positions in spectral space with 25 or fewer other pixels were considered to be nonland features and were removed, as shown in red in Figure 25. It can be seen in the illustration that the cumulus clouds on the right side and bottom of the scene were sucessfully removed. The stringency of screening can be controlled by the threshold of the VIS-IR repeats selected as the limit of cloud-free land features. In this summer daytime scene the cool water bodies of limited geographic extent and differing reflectances exhibited sufficient variability in spectral space to be classed as nonland by the cluster method (Figure 25).

Step 3 requires examination of the thermal image of the scene to determine if SCi is present. An affirmative answer comes from Figure 11.

Step 4 calls for determining the IR statistics of the scene after the visible clouds have been removed (Step 2). The selected upper threshold of IR=110 digital counts for SCi contaminated pixels was -2.8 standard deviations from the Ch. 2 mean.

Step 5 involves removal of SCi contamination, the coldest pixels remaining after cloud screening. This has been

done in Figure 26 where we show in red where pixels having IR digital counts of 110 or less were removed. As desired, the SCi affected area in the center of the scene and the area at about 1/3 of the distance from the top and 1/4 from the right edge come out. At the expense of also removing good data, more of the thin SCi fringe surrounding the affected areas could have been removed by using a higher threshold value.

Step 6 reports that the pixels that remain after screening are free of clouds and SCi. The uncontaminated pixels that remained after both screening procedures are shown in Figure 27. We note by comparison of Figures 25 and 26 that many of the same pixels were discarded by both screening procedures.

5.3 Considerations and limitations in editing

The example of screening discussed in section 5.2 illustrates some of the problems of editing satellite data for cloud and SCi contamination by the methods of section 5.1. The geographic area of south Texas and northeast Mexico used for the example created a complication because of the natural climatic gradient across it (Figures 10 and 11, and Ch. 2 trace in Figure 12). The right (east) edge of the area is at the Gulf of Mexico and this coastal region is cooler, with more available soil moisture, than the inland semiarid rangeland to the left. Average annual rainfall decreases about 1 cm/4.5 km from right to left across much of the 240 km shown.

The threshold IR digital value that was used in SCi screening in the above editing procedure was a compromise. Ideally more of the SCi fringes on the edges of the main SCi concentrations should have been removed. However, to do so by increasing the threshold digital value, would have removed more valid data in the cool coastal region.

Complications from strong surface temperature gradients are possible in other settings, such as from change of temperature with elevation in mountainous areas. In many cases, the problem could be helped by limiting the size of the area that is tackled per analysis.

5.4 Conclusions

Pixels affected by clouds and/or SCi can be successfully removed from satellite data by the two-stage procedure that uses reflective and emissive wave bands, as discussed in this chapter. The procedure first removes clouds most readily detectable in the VIS bands and then removes SCi from the data that remain.

The size of area, as a process entity, that can best be corrected depends on the thermal contrast in the image. A large area can be processed as a unit if strong surface temperature gradients do not extend across it. Situations such as marked coastal climatic gradients or large elevation differences, with attendant temperature variations, may best be processed in 16,000 km^2 or smaller segments.

6.0 LITERATURE CITED

Nixon, P. R. 1982. Selective post freeze sugarcane harvesting using satellite data. IN Texas sugarcane growers guide. Texas Agricultural Experiment Station, College Station, Texas (In Press).

Tarpley, J. D. 1979. Estimating incident solar radiation at the surface from geostationary satellite data. Jour. Applied Meteor. 18:1172-1181.

Wiegand, C. L., P. R. Nixon, H. W. Gausman, L. N. Namken, R. W. Leamer, and A. J. Richardson. 1981. Plant cover, soil temperature, freeze, water stress, and evapotranspiration conditions. Type III Final Report prepared for Goodard Space Flight Center. U.S. Department of Agriculture, Weslaco, TX 78596.

Figure 1. Location of daytime HCMM data is spectral space, reflective (VIS) vs. emissive (IR). These 3 Jul 78 data are from a 100,000 pixel Lower Rio Grande Valley of Texas test area. See text for identification of symbols.

Figure 2. The 3 Jul 78 HCMM daytime data in spectral space, after removal of cloud contaminated data using a modified cluster screening method. See text for identification of symbols.

Figure 3. Distribution of reflective (VIS) data in a 100,000 pixel Lower Grande Valley of Texas test area on 3 Jul 78.

Figure 4. Distribution of remaining reflective (VIS) data in a 100,00 pixel Lower Rio Grande Valley of Texas test area on 3 Jul 78 after cloud screening by a modified cluster method.

Figure 5. Distribution of emissive (IR) data in a 100,000 pixel Lower Rio Grande Valley of Texas test area on 3 Jul 78.

Figure 6. Distribution of remaining emissive (IR) data in a 100,000 pixel Lower Rio Grande Valley of Texas test area on 3 Jul 78 after cloud screening by a modified cluster method.

Figure 7. Location of 100,000-pixel Lower Rio Grande Valley of Texas test area in spectral space on 3 Jul 78 (upper illustration). Location of selected cloud, land and water features in spectral space on 3 Jul 78 (lower illustration). Consult Table 1 for identification of letters.

Figure 8. Location of 100,000-pixel Lower Rio Grande Valley of Texas test area in spectral space on 15 Aug 78 (upper illustration). Location of selected cloud, land and water features in spectral space on 15 Aug 78 (lower illustration). Consult Table 1 for identification of letters.

Figure 9. Location of 100,000-pixel Lower Rio Grande Valley of Texas test area is spectral space on 7 Feb 79 (upper illustration). Location of selected cloud, land and water features in spectral space on 7 Feb 79 (lower illustration). Consult Table 1 for identification of letters.

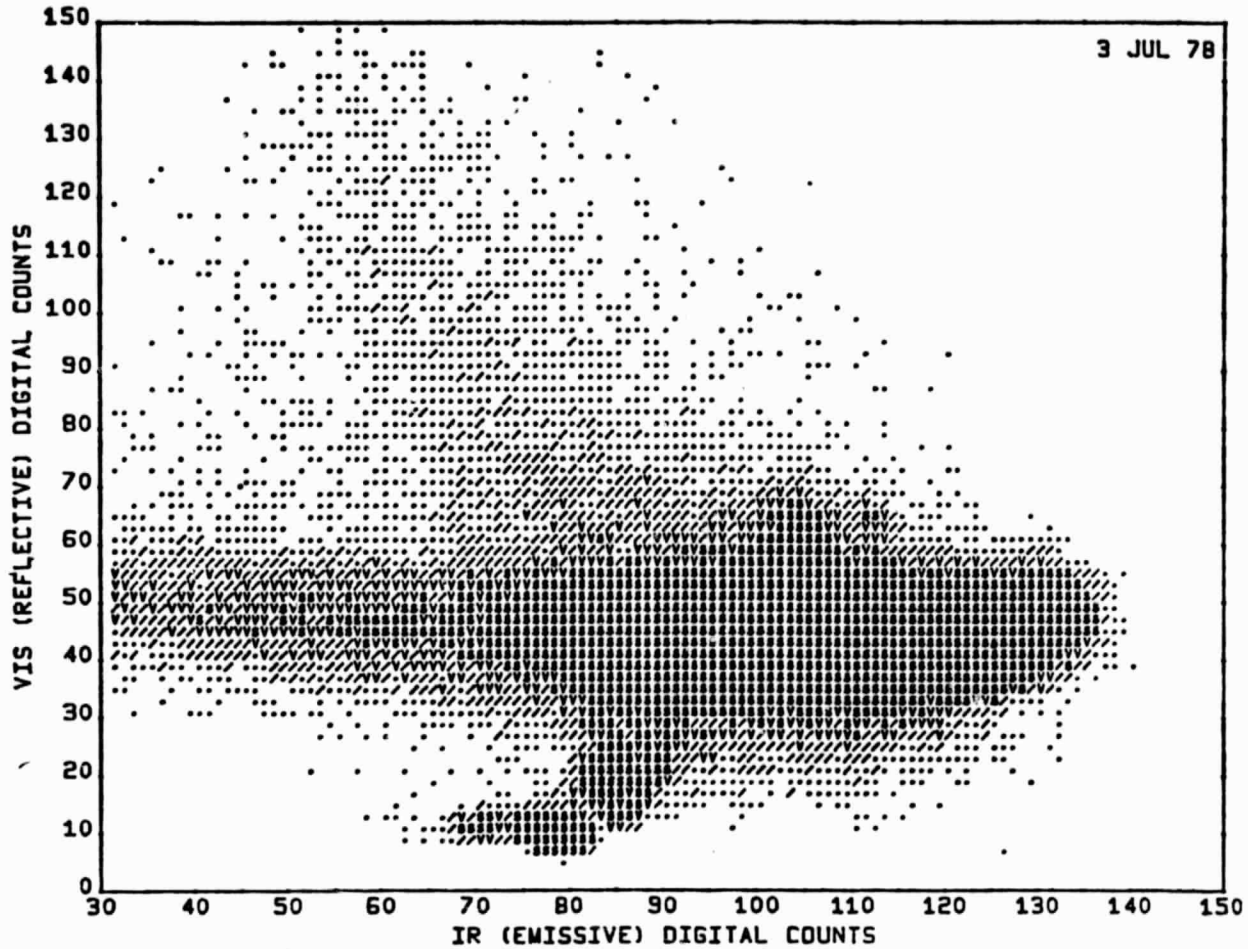

Figure 1. Location of daytime HCMM data is spectral space, reflective (VIS) vs. emissive (IR). These 3 Jul 78 data are from a 100,000 pixel Lower Rio Grande Valley of Texas test area. See text for identification of symbols.

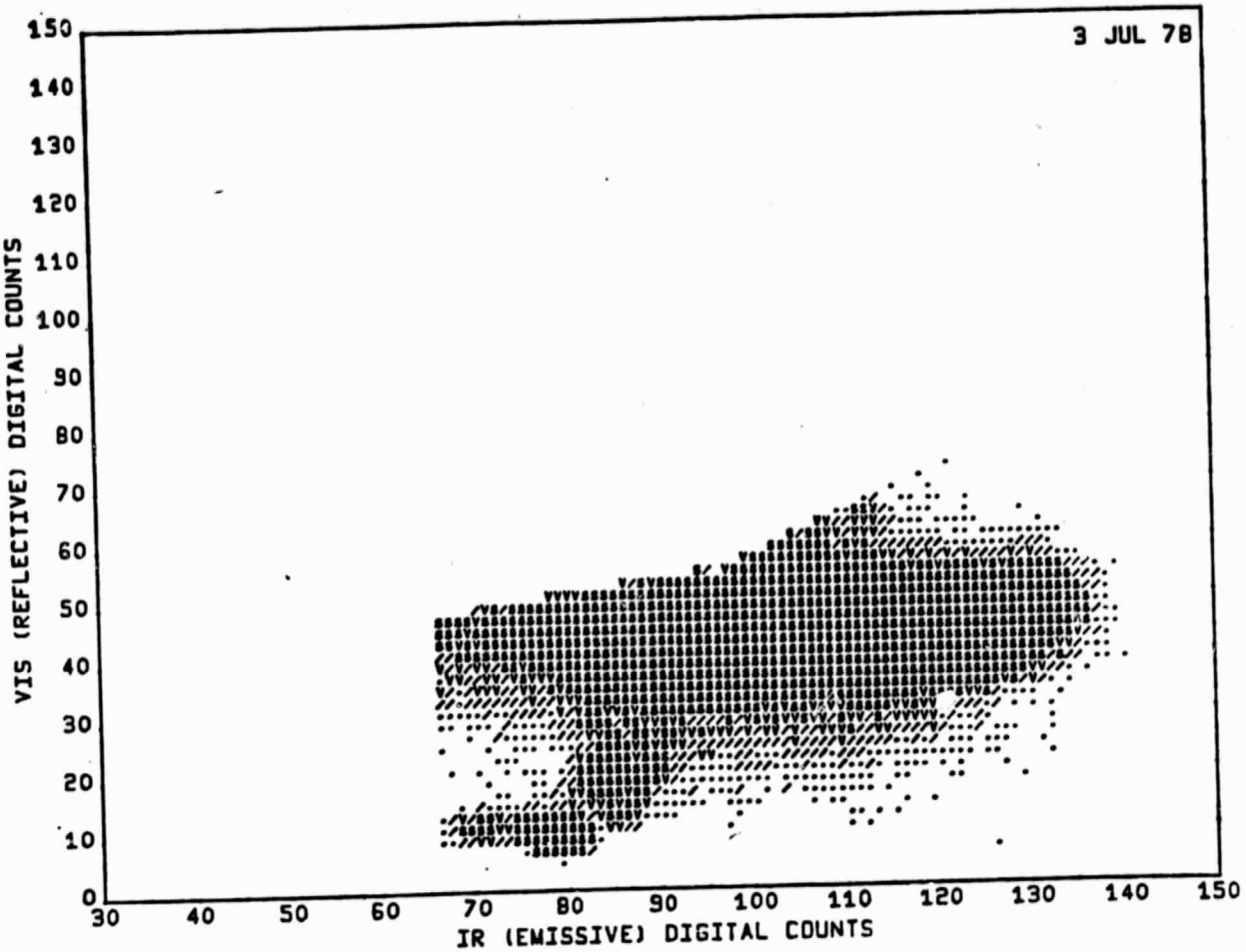

Figure 2. The 3 Jul 78 HCMM daytime data in spectral space, after removal of cloud contaminated data using a modified cluster screening method. See text for identification of symbols.

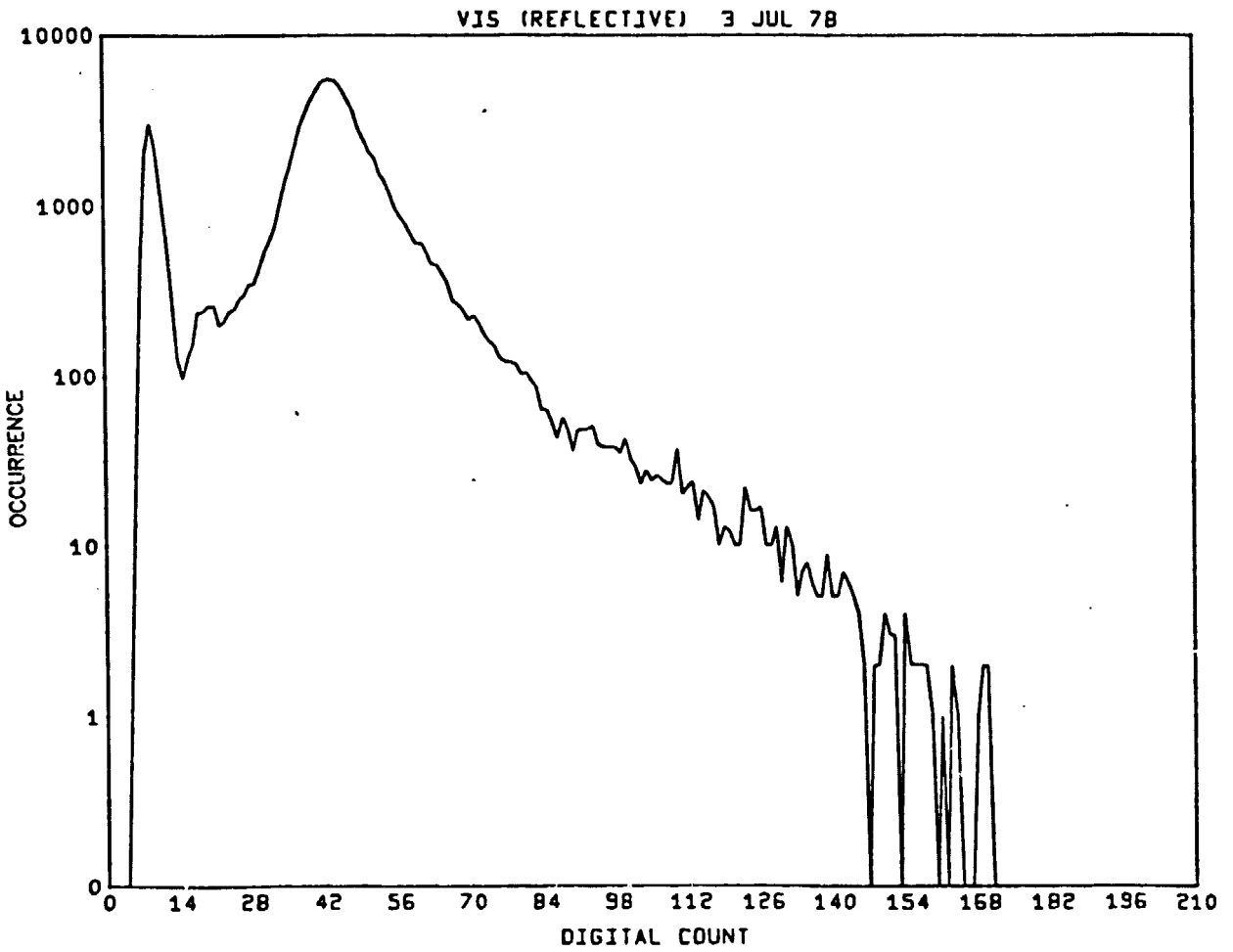

Figure 3. Distribution of reflective (VIS) data in a 100,000 pixel Lower Rio Grande Valley of Texas test area on 3 Jul 78.

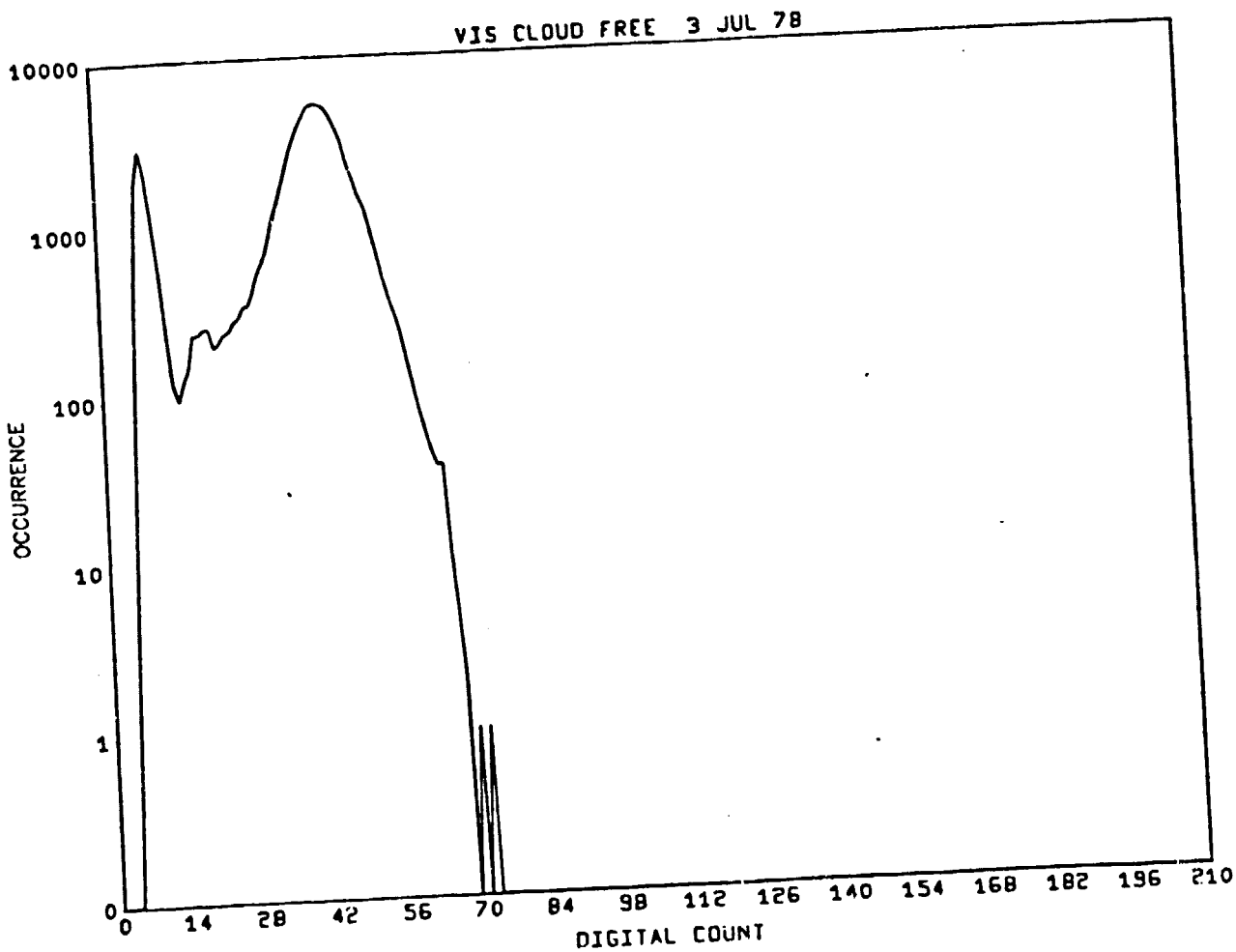

Figure 4. Distribution of remaining reflective (VIS) data in a 100,000 pixel Lower Rio Grande Valley of Texas test area on 3 Jul 78 after cloud screening by a modified cluster method.

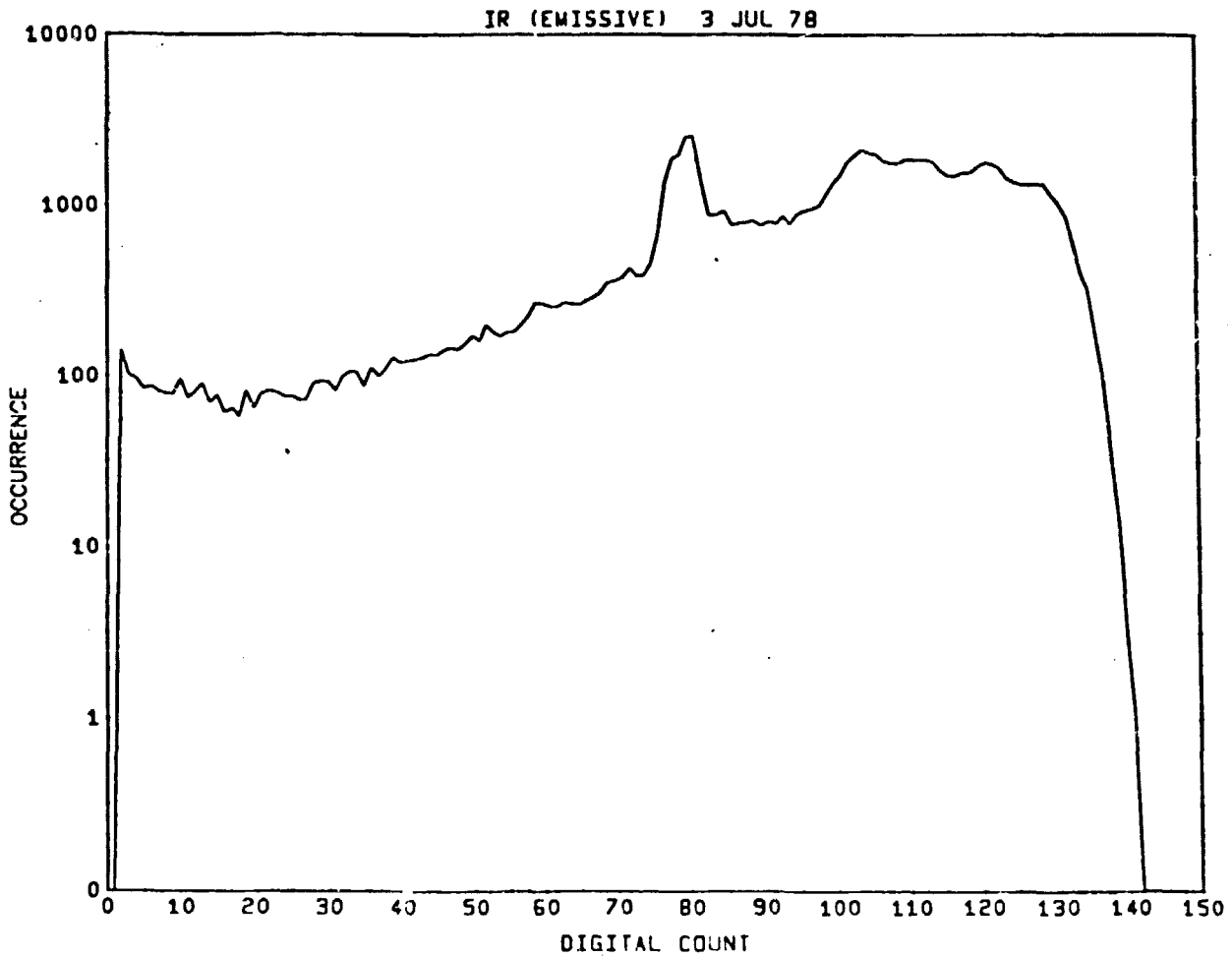

Figure 5. Distribution of emissive (IR) data in a 100,000 pixel Lower Rio Grande Valley of Texas Test area on 3 Jul 78.

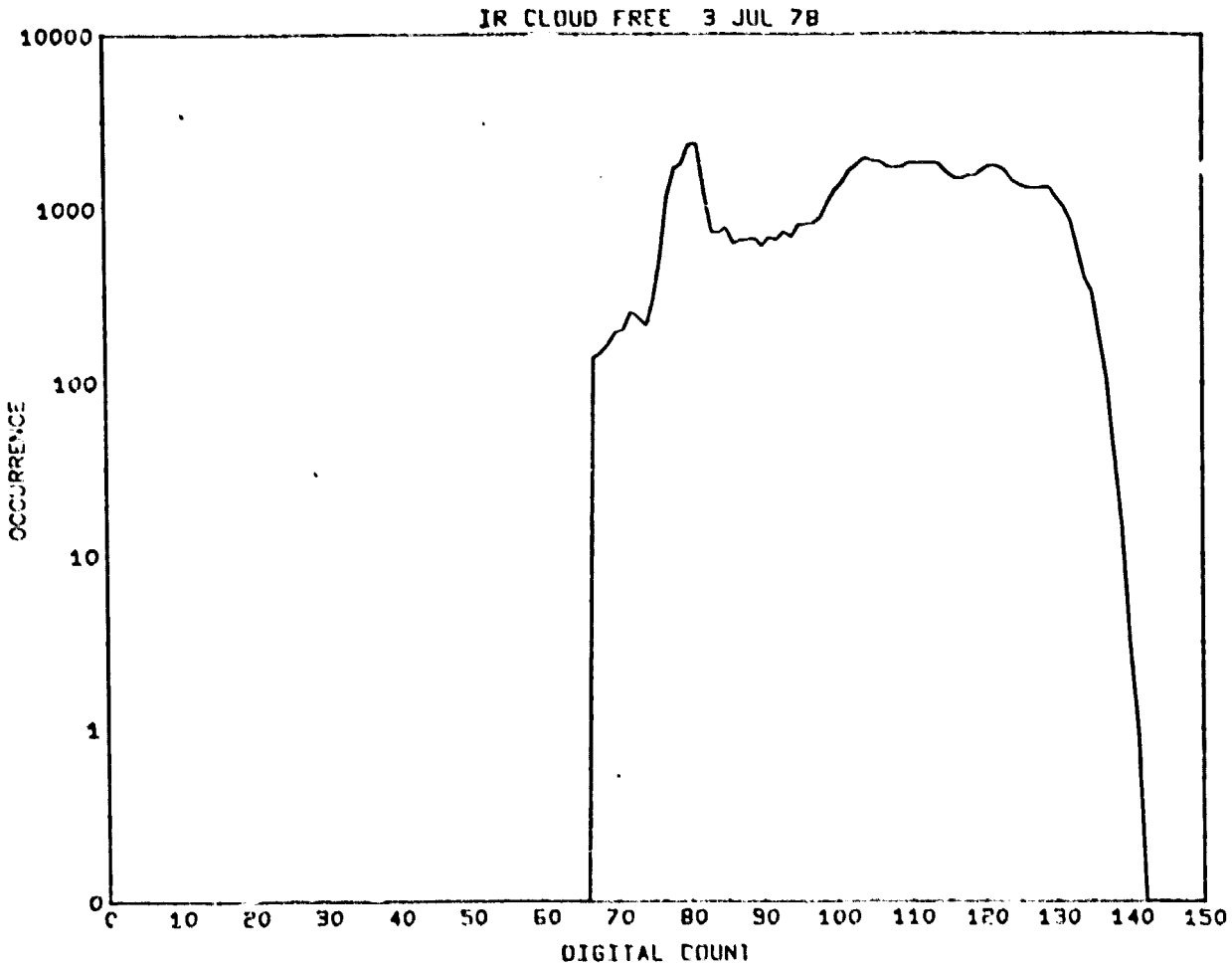

Figure 6. Distribution of remaining emissive (IR) data in a 100,000 pixel Lower Rio Grande Valley of Texas test area on 3 Jul 78 after cloud screening by a modified cluster method.

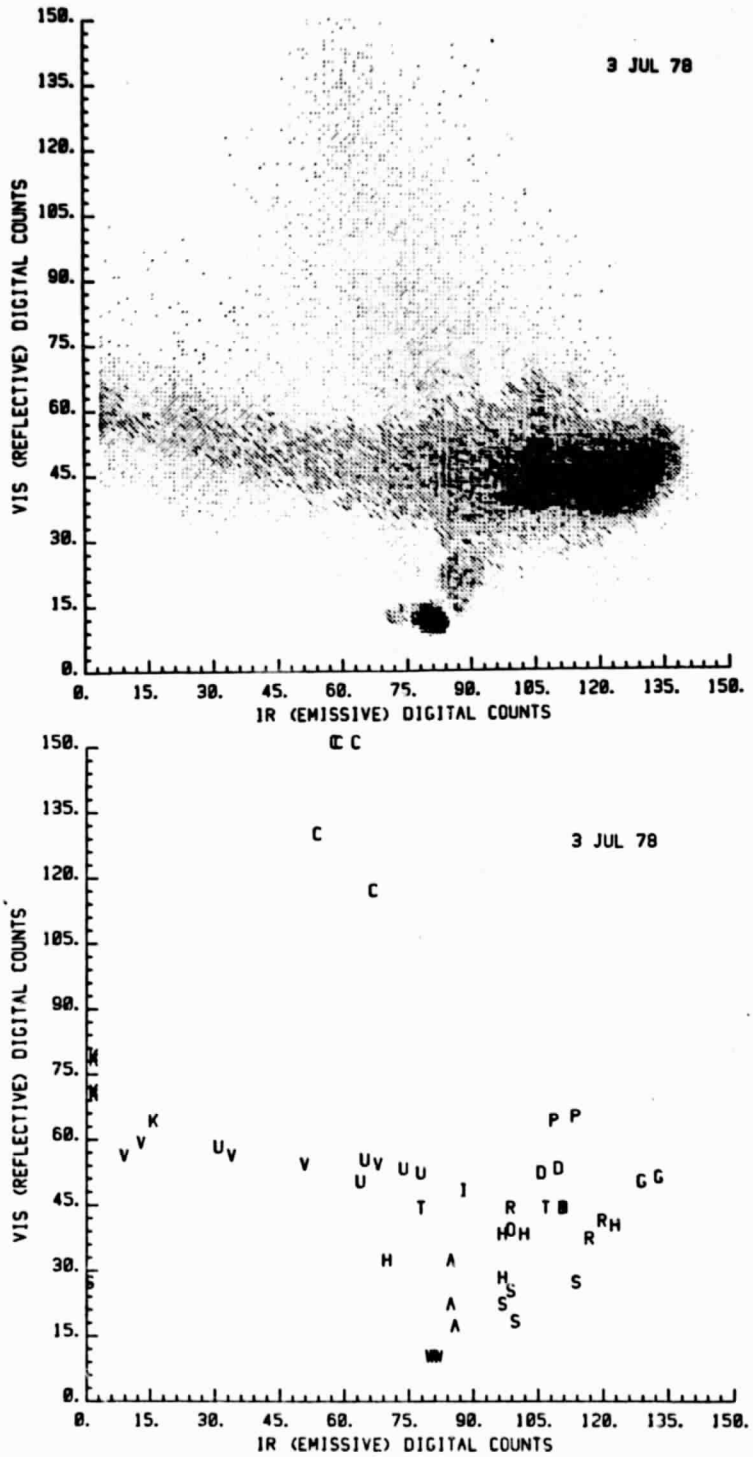

Figure 7. Location of 100,000-pixel Lower Rio Grande Valley of Texas test area in spectral space on 3 Jul 78 (upper illustration). Location of selected cloud, land and water features in spectral space on 3 Jul 78 (lower illustration). Consult Table 1 for identification of letters.

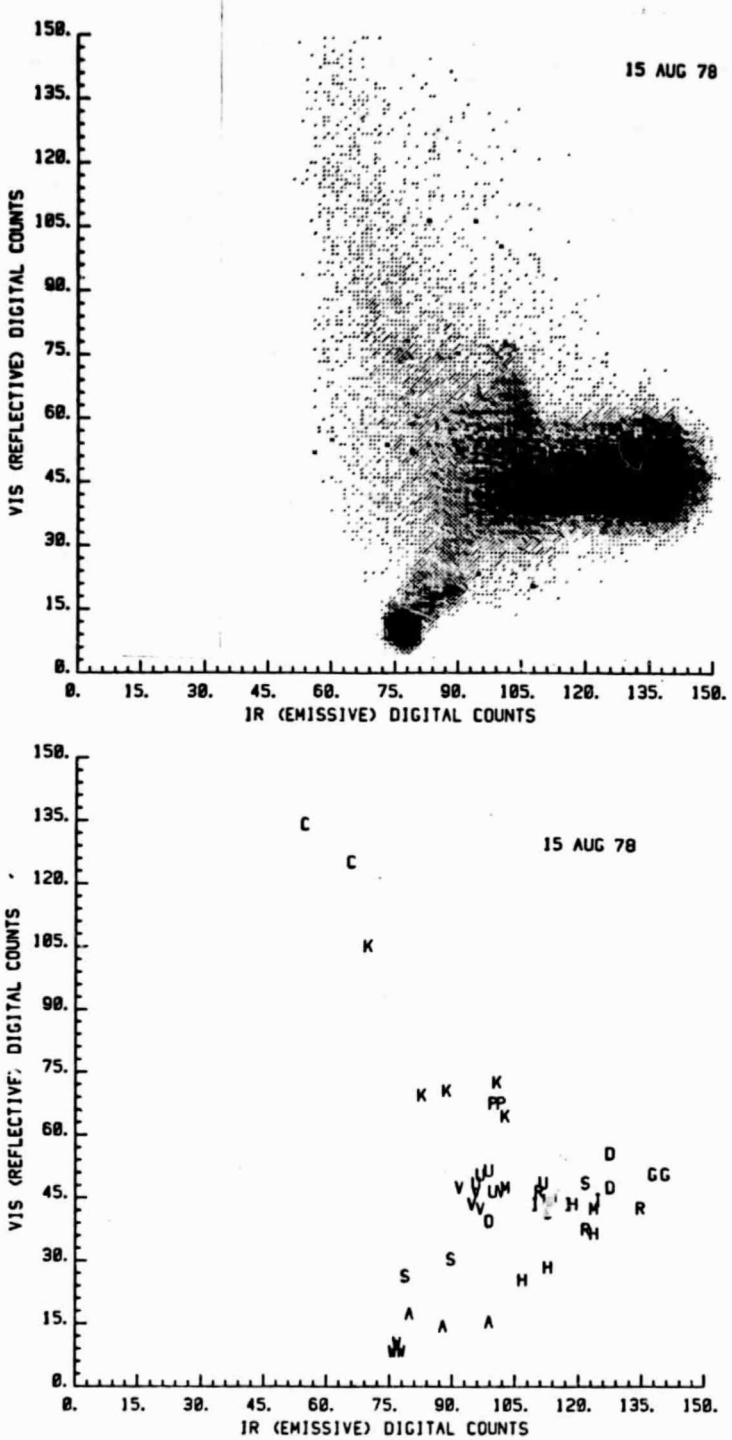

Figure 8. Location of 100,000-pixel Lower Rio Grande Valley of Texas test area in spectral space on 15 Aug 78 (upper illustration). Location of selected cloud, land and water features in spectral space on 15 Aug 78 (lower illustration). Consult Table 1 for identification of letters.

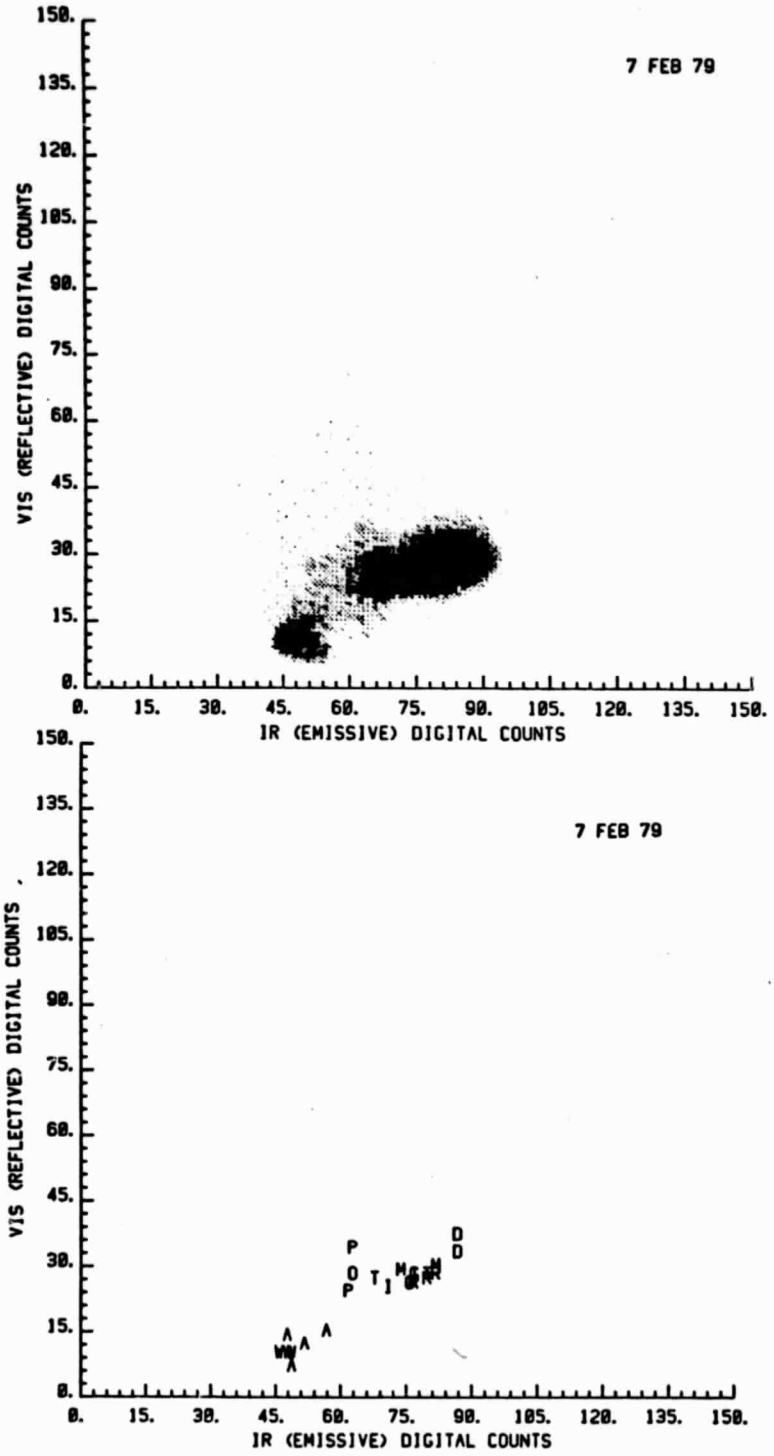

Figure 9. Location of 100,000-pixel Lower Rio Grande Valley of Texas test area in spectral space on 7 Feb 79 (upper illustration). Location of selected land and water features in spectral space on 7 Feb 79 (lower illustration). Consult Table 1 for identification of letters.

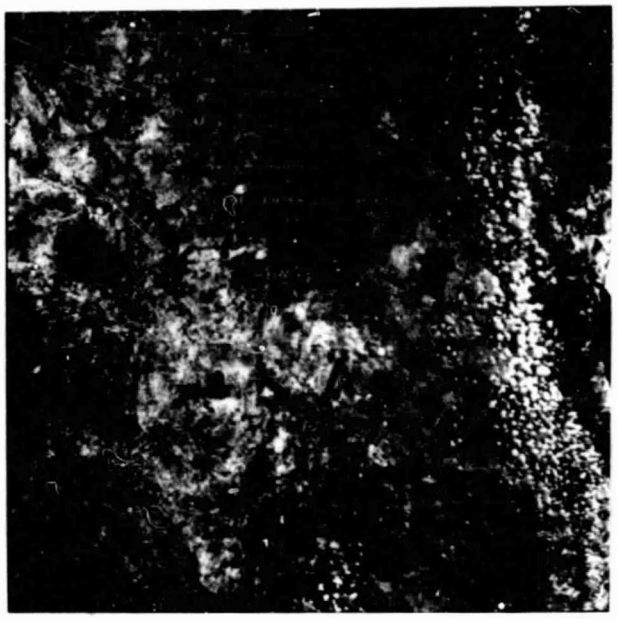

Figure 10. HCMM daytime VIS scene of south Texas and northeast Mexico on 15 Aug 78. Scattered cumulus clouds parallel the coast (right and lower center).

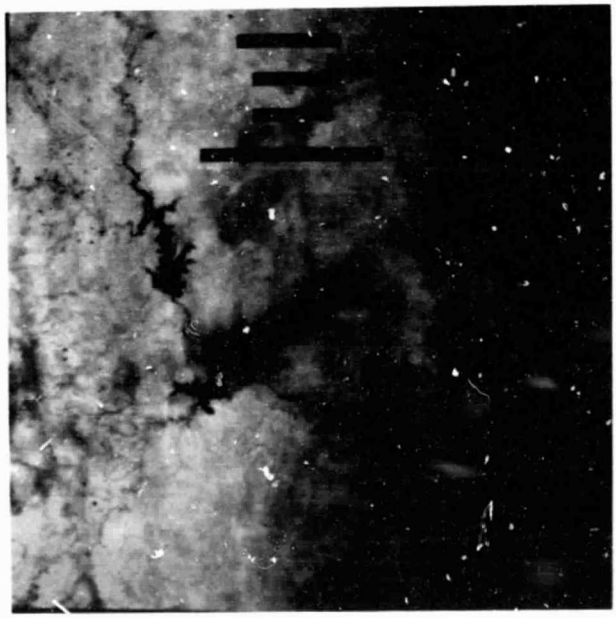

Figure 11. HCMM daytime IR scene of south Texas and northeast Mexico on 15 Aug 78. In addition to clouds an atmospheric absorbing-emitting layer (wispy dark areas center to upper right) is apparent in this IR band scene.

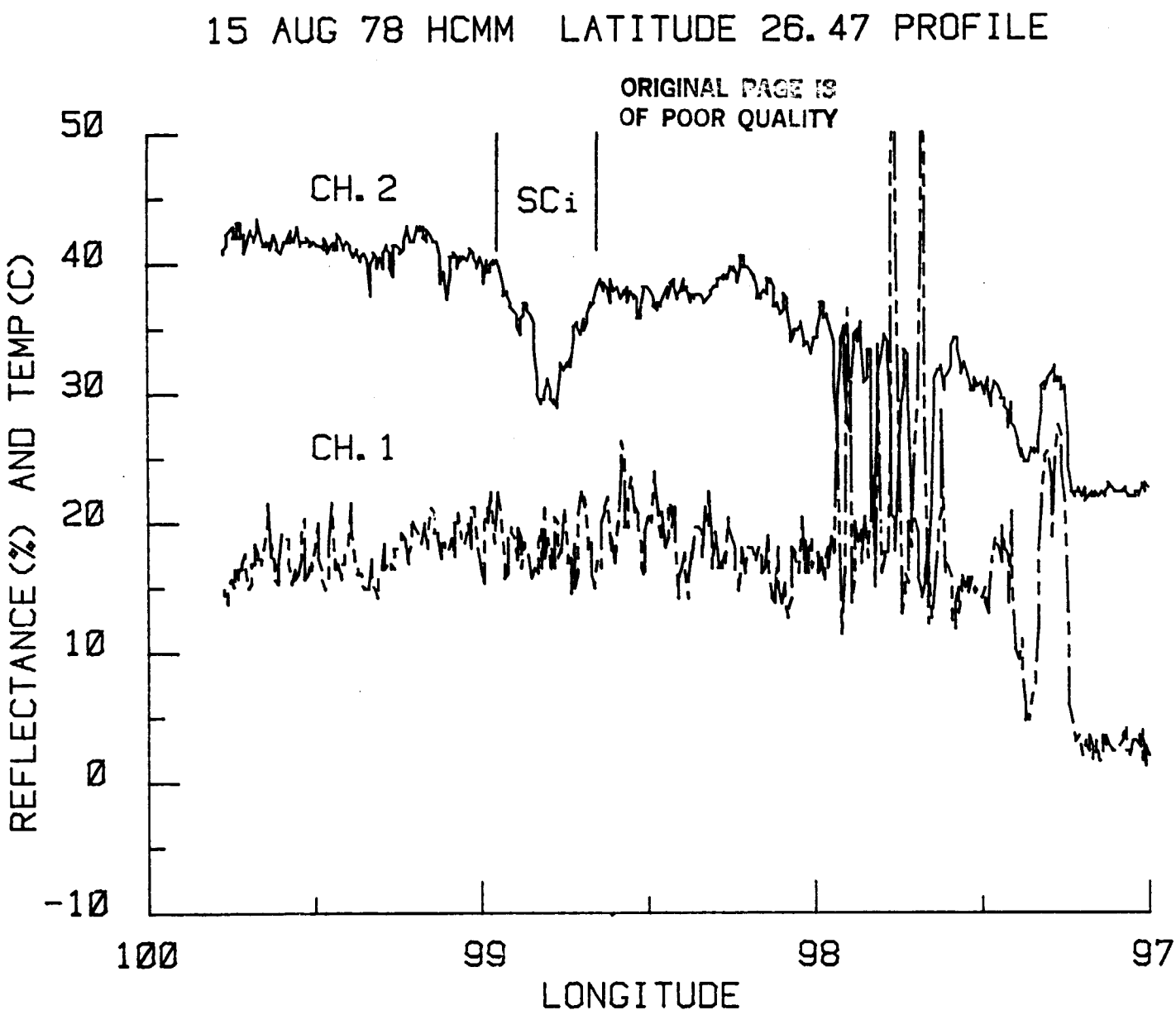

Figure 12. Reflected (VIS, Ch. 1) and emitted (IR, Ch. 2) radiation across the HCMM south Texas test area on latitude 26.47 N., 15 Aug 78.

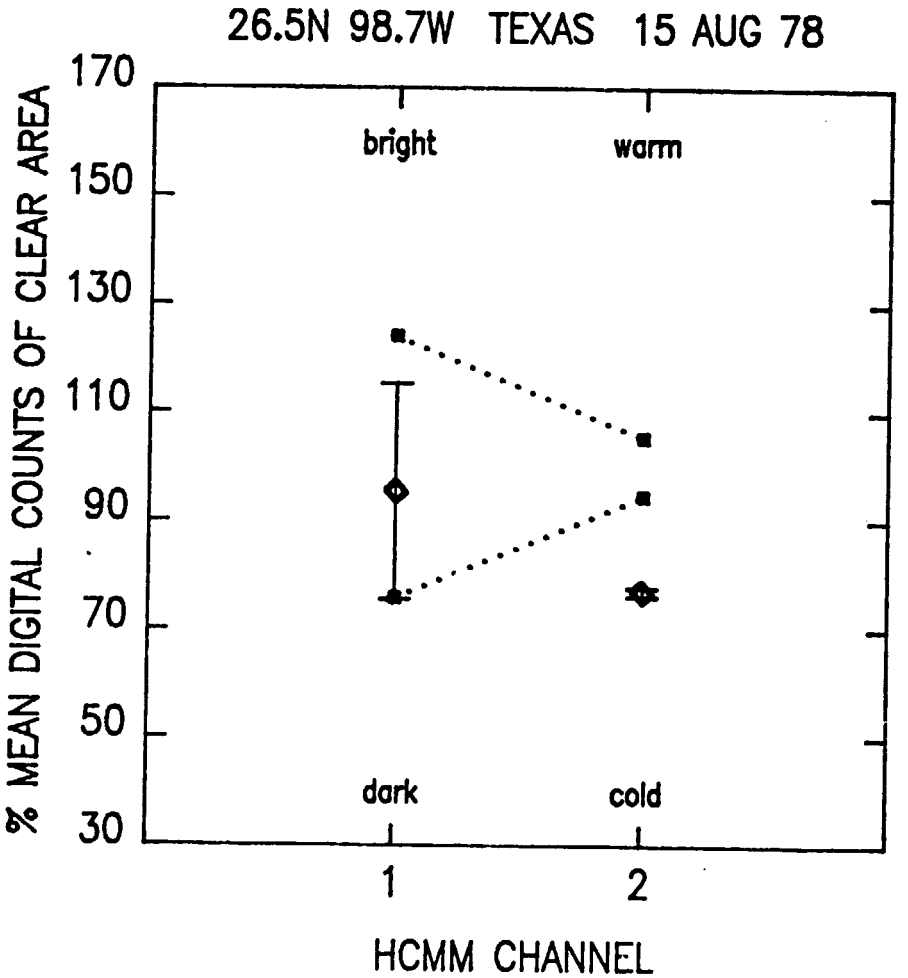

Figure 13, HCMM spectral statistics of SCi and adjacent clear area in south Texas, 15 Aug 78. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

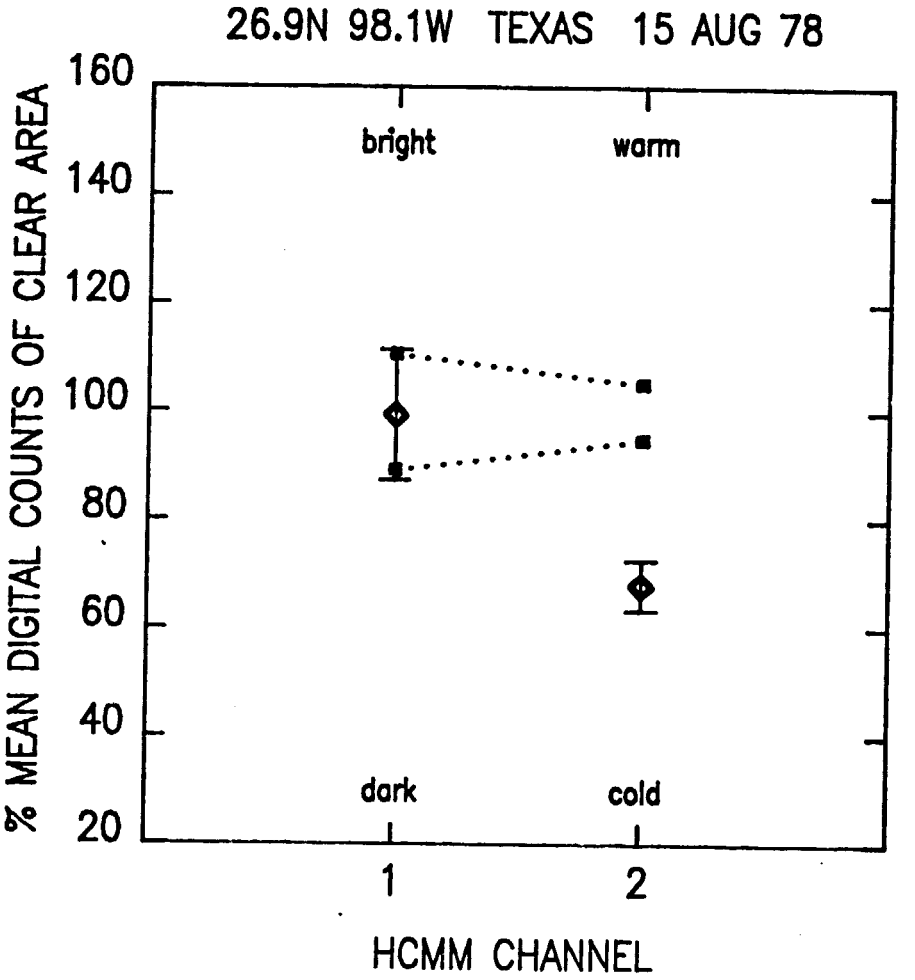

Figure 14. HCMM spectral statistics of SCi and adjacent clear area in coastal Texas, 15 Aug 78. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

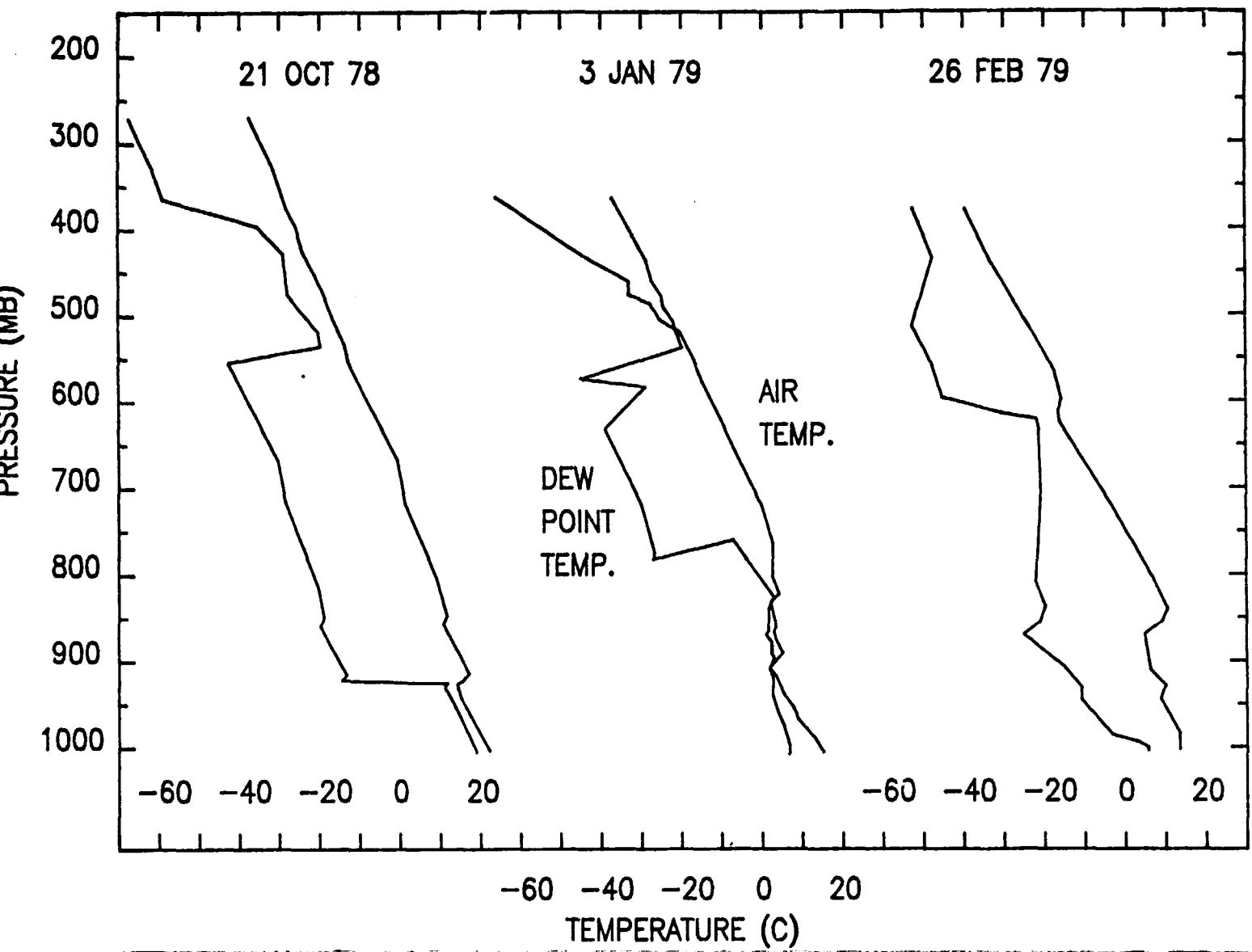

Figure 15. Radiosonde observations at Brownsville, Texas on the dates shown. These atmospheric soundings were made at 0600 hrs, except for 3 Jan 79 when the sounding was made at 1800 hrs.

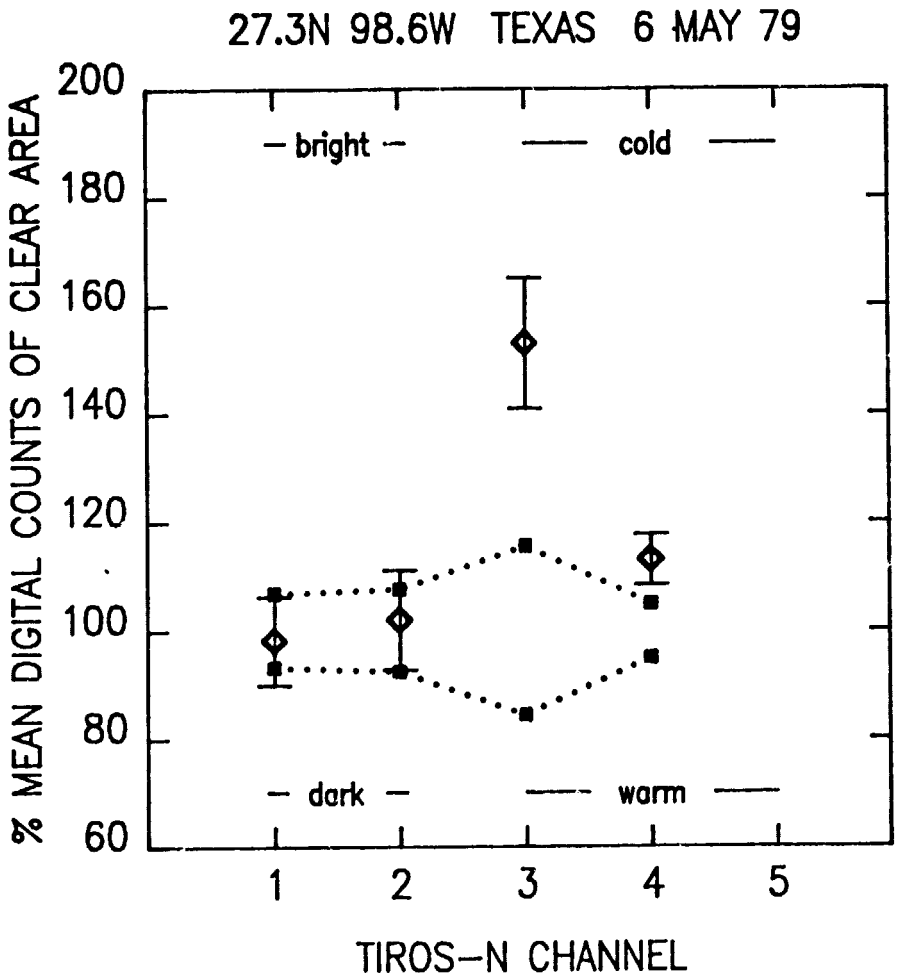

Figure 16. TIROS-N spectral statistics of SCi and adjacent clear area in south Texas, 6 May 79. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

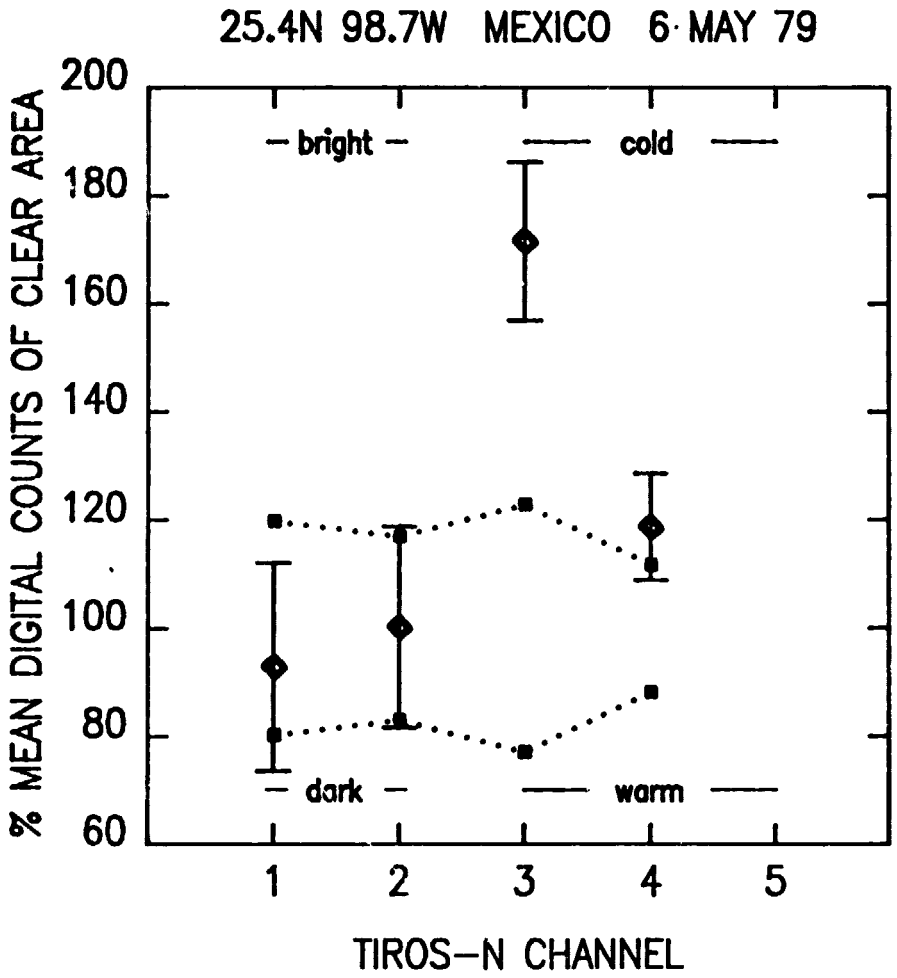

Figure 17. TIROS-N spectral statistics of SCi and adjacent clear area in northeast Mexico, 6 May 79. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

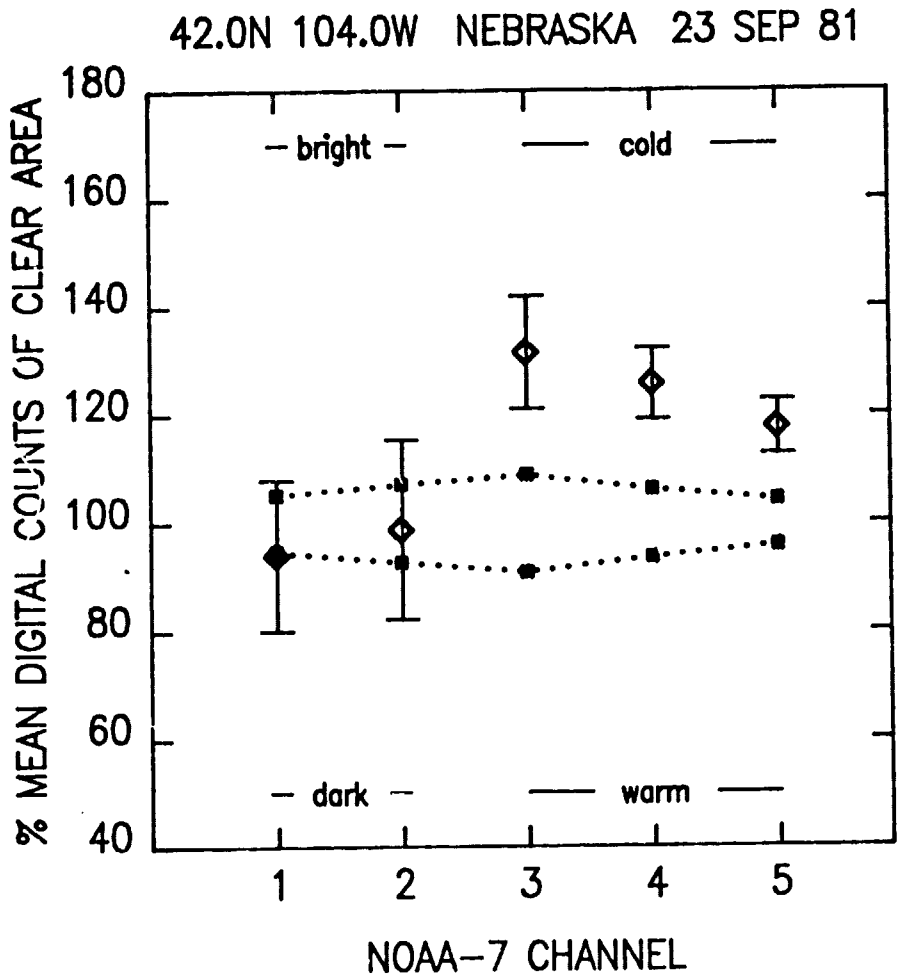

Figure 18, NOAA-7 spectral statistics of SCi and adjacent clear area in Nebraska, 23 Sep 81. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

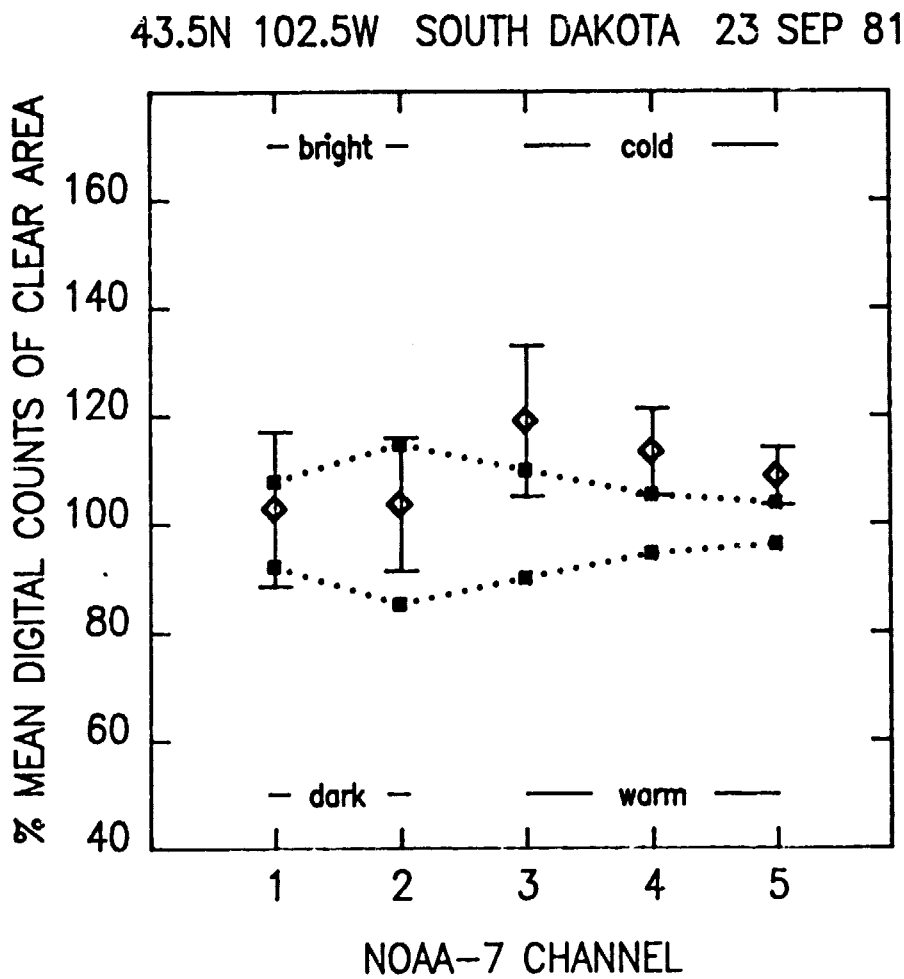

Figure 19. NOAA-7 spectral statistics of SCi and adjacent clear area in South Dakota, 23 Sep 81. Means and ±2 standard deviations are shown normalized to clear are means (100 percent).

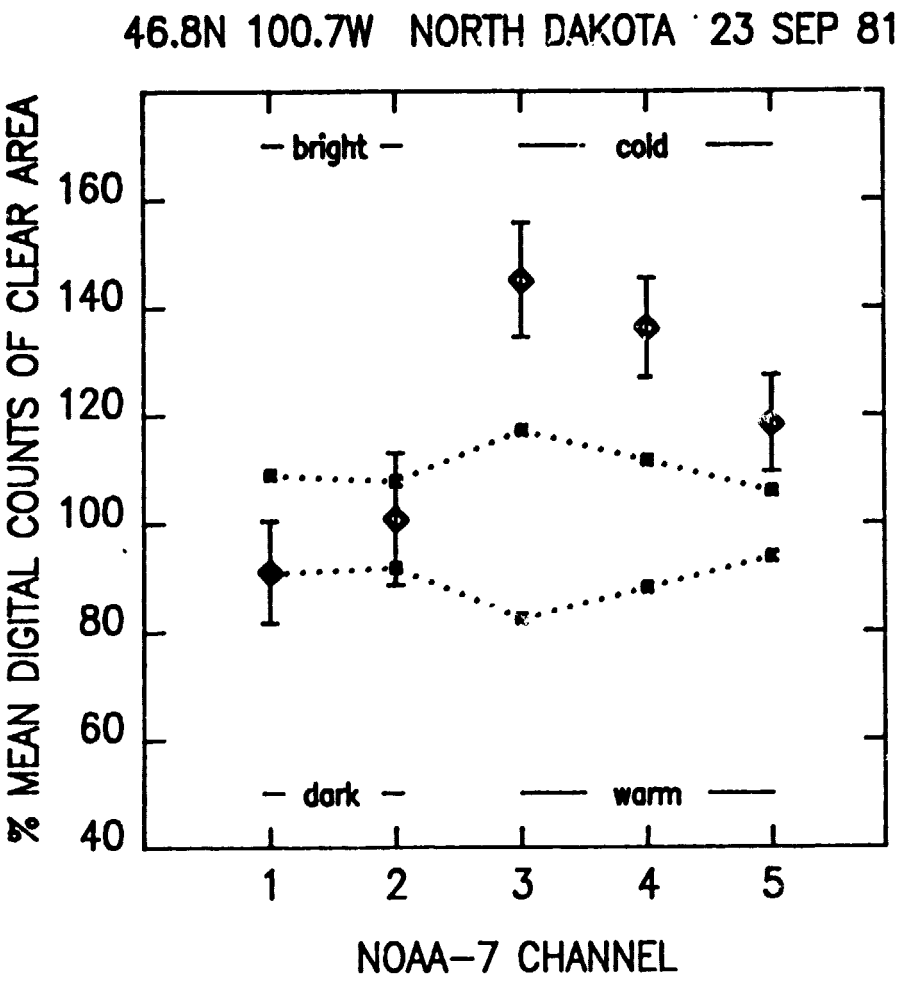

Figure 20. NOAA-7 spectral statistics of SCi and adjacent clear area in North Dakota, 23 Sep 81. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

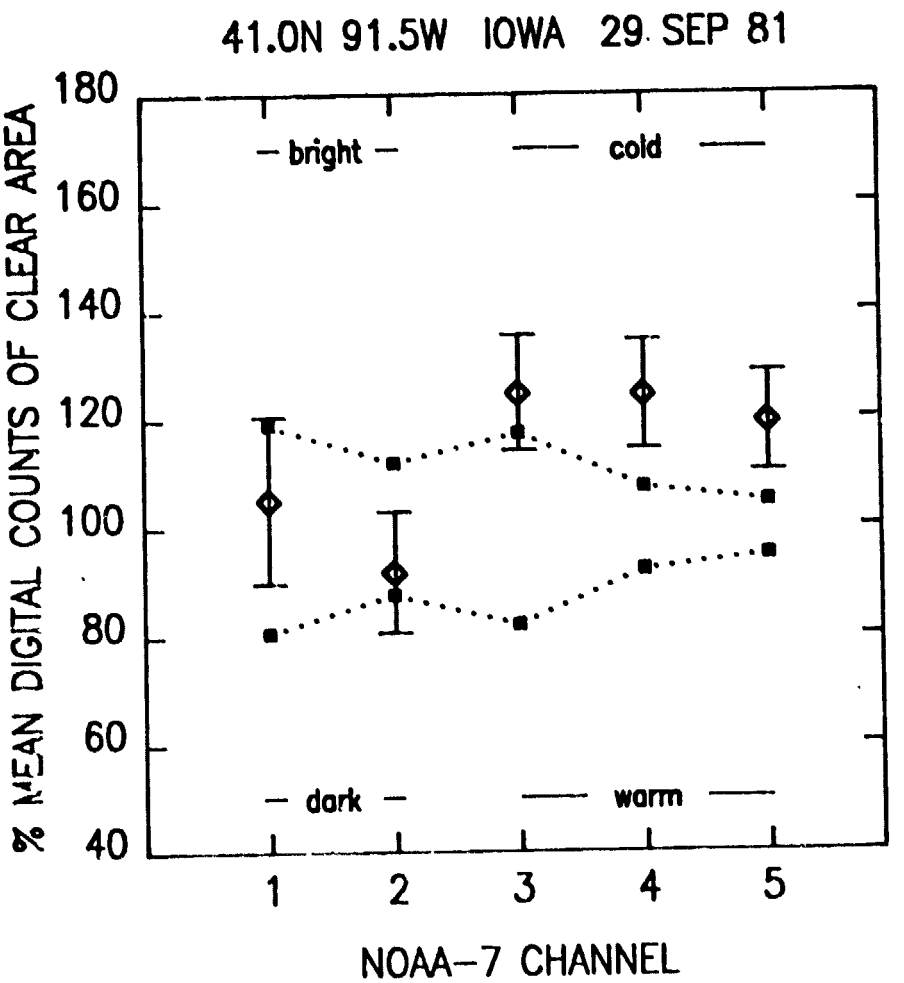

Figure 21. NOAA-7 spectral statistics of SCi in Iowa and adjacent clear area, 29 Sep 81. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

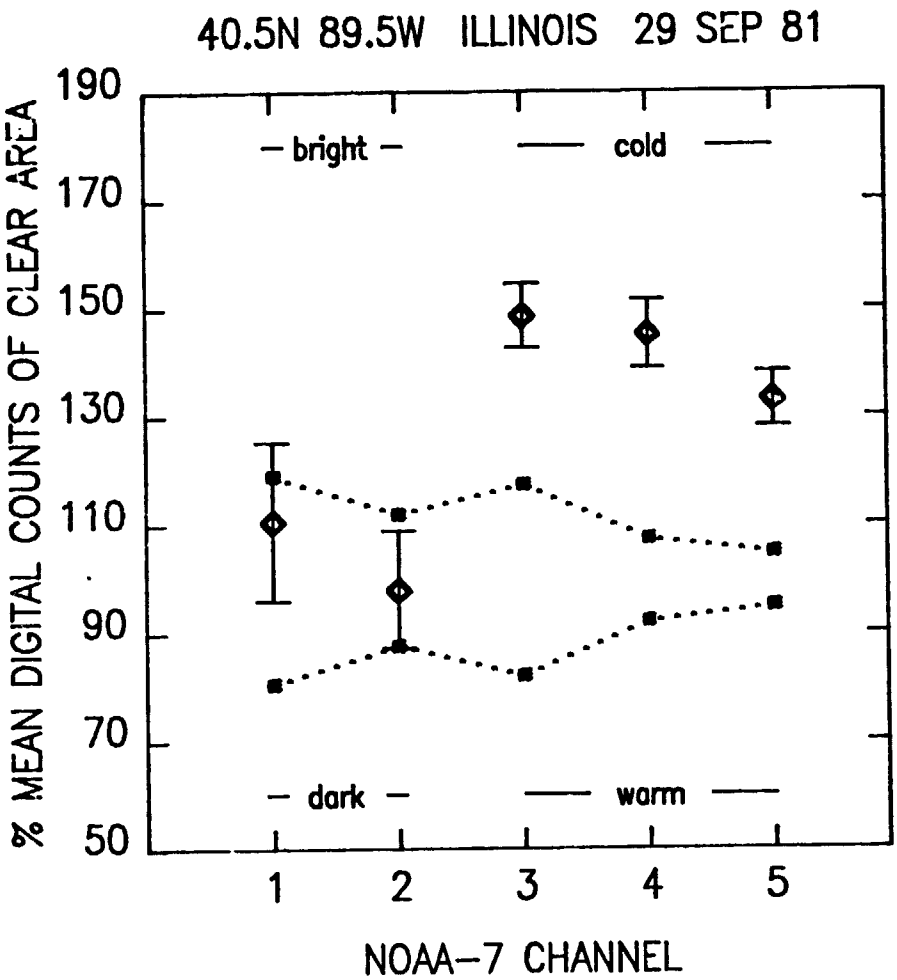

Figure 22. NOAA-7 spectral statistics of SCi in Illinois and adjacent clear area, 29 Sep 81. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

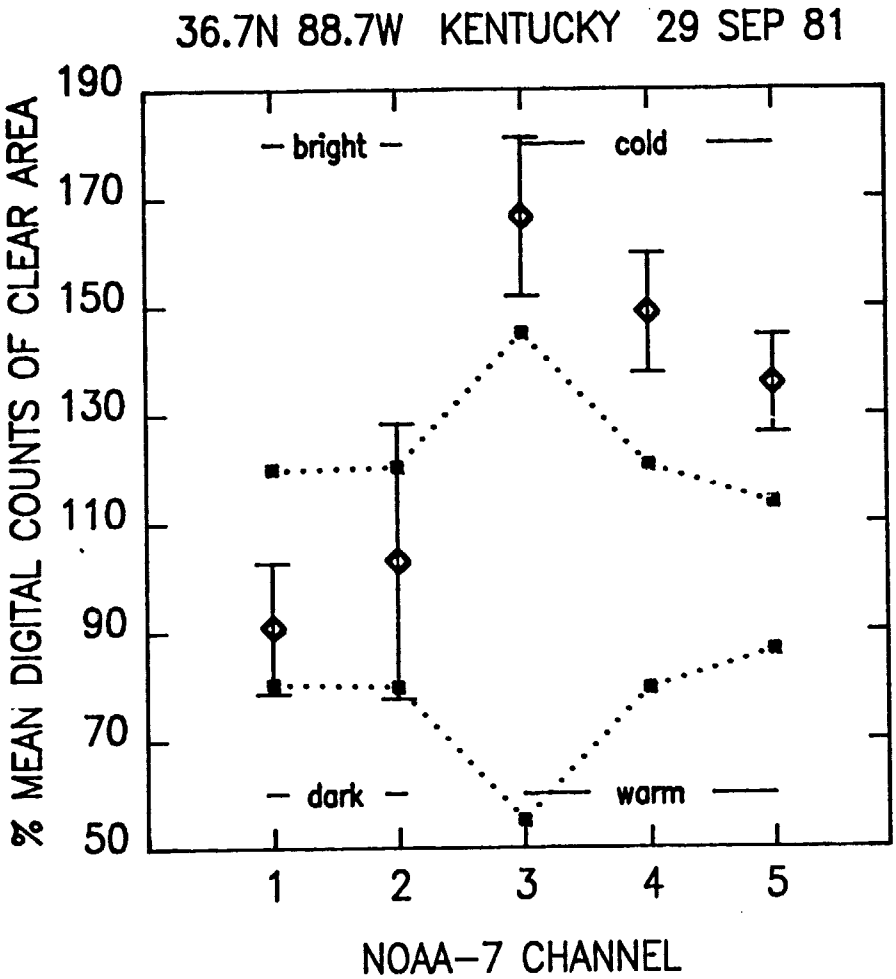

Figure 23. NOAA-7 spectral statistics of SCi and adjacent clear area in Kentucky, 29, Sep 81. Means and ±2 standard deviations are shown normalized to clear area means (100 percent).

ORIGINAL PAGE
COLOR PHOTOGRAPH

Figure 24. Location of a 262,144-pixel south Texas and northeast Mexico region in spectral space. HCMM data of 15 Aug 78 where Ch. 1 is refective VIS and Ch. 2 is emissive IR.

Figure 25. Locations shown in red of pixels removed by cluster screening.

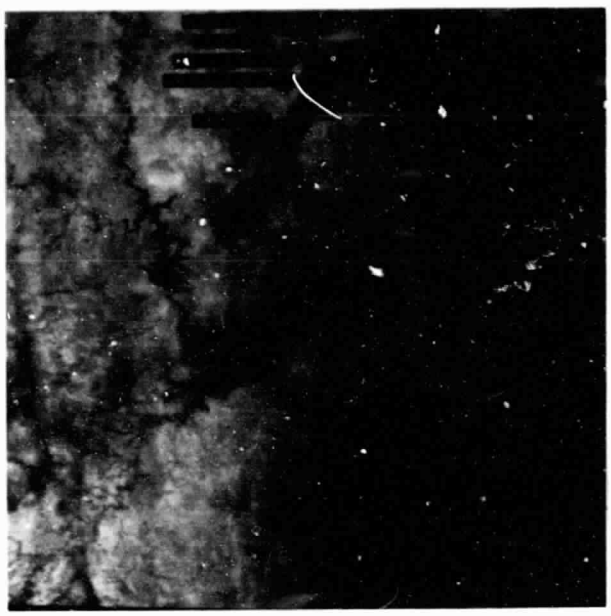

Figure 26. Locations shown in red of pixels removed by SCi screening based on IR digital values of 110 or less.

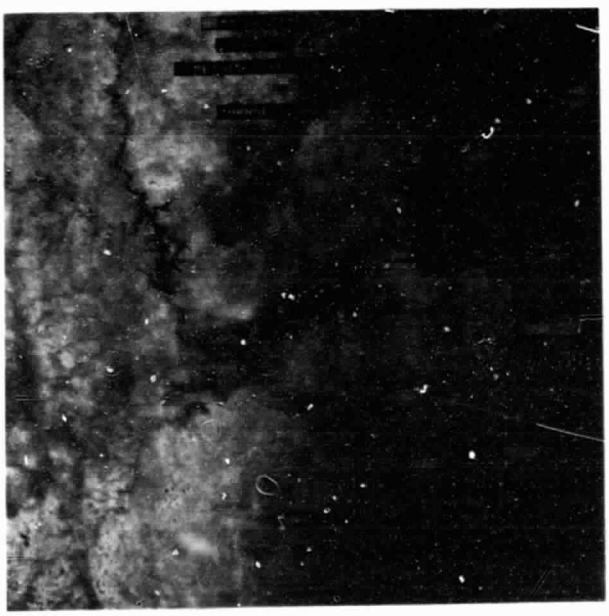

Figure 27. Location shown in red of pixels removed by the combination of cloud and SCi screening techniques.

Table 1. Sources of polar-orbiting satellite data used in this report, including location of center of scene, prevailing sun zenith angle and reference to figure number in which the data are presented.

Figure	Satellite	Date	Overpass Time (CST)	Location Lat	Location Long	Zenith Angle
1-7	HCMM	3 Jul 78	14:06	26.7	98.6	20.74
8	HCMM	15 Aug 78	14:08	26.7	98.6	23.64
9	HCMM	7 Feb 79	13:42	26.7	98.6	44.89
10-11	HCMM	15 Aug 78	14:04	26.7	98.6	23.64
12	HCMM	15 Aug 78	14:04	26.5	98.4	23.68
13	HCMM	15 Aug 78	14:04	26.5	98.7	23.47
14	HCMM	15 Aug 78	14:04	26.9	98.1	24.07
16	TIROS-N	6 May 79	14:37	27.3	98.6	35.95
17	TIROS-N	6 May 79	14:37	25.4	98.7	30.99
18	NOAA-7	23 Sep 81	14:10	42.0	104.0	43.43
19	NOAA-7	23 Sep 81	14:16	43.5	102.5	45.34
20	NOAA-7	23 Sep 81	14:23	46.8	100.7	49.56
21	NOAA-7	29 Sep 81	14:29	41.0	91.5	56.37
22	NOAA-7	29 Sep 81	14:37	40.5	89.5	58.44
23	NOAA-7	29 Sep 81	14:40	36.7	88.7	59.70
24-27	HCMM	15 Aug 78	14:04	26.7	98.6	23.64

HCMM
Ch 1 = 0.55 to 1.10 μm, Ch 2 = 10.5 to 12.5 μm

TIROS-N
Ch 1 = 0.55 to 0.90 μm, Ch 2 = 0.725 to 1.10 μm, Ch 3 = 3.55 to 3.93 μm, Ch 4 = 10.5 to 11.5 μm

NOAA-7
Ch 1 = 0.58 to 0.68 μm, Ch 2 = 0.725 to 1.10 μm, Ch 3 = 3.55 to 3.93 μm, Ch 4 = 10.3 to 11.3 μm, Ch 5 = 11.5 to 12.5 μm

Table 2. Identification of selected cloud, land and water features that are shown by symbols in Figures 7, 8, and 9.

Symbol	Number of sample sites	Number of pixels per sample	Site ident., if applicable	Identification
A	4	1	111,621,821,931	Inland water bodies
C	5	1		Coastal clouds
K	5	1		Inland clouds
D	2	5x5	421,422	Dryland sorghum
G	2	3x3	121,213	Buffelgrass (non-irrigated)
I	3	9x9	533,631,633	Mixed irrigated crops
M	2	3x3	531,731	Cities, Harlingen McAllen
O	1	3x3	701	Oak trees, extensive grove
P	3	1	801	Padre Island dune sand
R	3	9x9	214,312,501	Rangeland
S	5	1		Coastal cloud shadows
H	5	1		Inland cloud shadows
T	2	3x3	424,632	Irrigated citrus
U	5	3x3		Inland subvisible cirrus
V	5	3x3		Coastal subvisible cirrus
W	1	3x3	901	Gulf of Mexico water